PUBLICATIONS DE L'INSTITUT DE STATISTIQUE
DE L'UNIVERSITÉ DE PARIS

MÉMOIRES ET CONFÉRENCES SUR LE CALCUL DES PROBABILITÉS,
LA STATISTIQUE THÉORIQUE ET APPLIQUÉE, L'ÉCONOMÉTRIE

R. FERON

SUR LES TABLEAUX DE CORRÉLATION DONT LES MARGES SONT DONNÉES

R. FÉRON

SUR LES DISTRIBUTIONS DE PROBABILITÉ DE LAURENT SCHWARTZ

M. GIRAULT

PRODUIT DE FONCTIONS CARACTÉRISTIQUES

S. K. NASR

DES POSITIONS TYPIQUES D'UNE VARIABLE ALÉATOIRE

A. WINTNER

DES DISTRIBUTIONS SYMÉTRIQUES A FONCTIONS CARACTÉRISTIQUES CONVEXES

ANALYSE D'ARTICLES

VOL. V - FASCICULE 1 - 1956

PARIS
11, Rue Pierre Curie

Toute la correspondance relative aux publications
doit être envoyée à l'adresse :

INSTITUT DE STATISTIQUE DE L'UNIVERSITE DE PARIS
INSTITUT HENRI POINCARÉ - 11, Rue Pierre Curie - PARIS (Ve)

Les manuscrits doivent être envoyés à *M. Daniel DUGUÉ*,
à l'adresse précédente.

Abonnements : Pour la France 1.700 francs français
Pour l'Etranger 2.000 francs
Vente au numéro : *(fascicule de 50 pages environ)*
Pour la France 500 francs
Pour l'Etranger 600 francs

PUBLICATIONS DE L'INSTITUT DE STATISTIQUE DE L'UNIVERSITÉ DE PARIS

MÉMOIRES ET CONFÉRENCES SUR LE CALCUL DES PROBABILITÉS, LA STATISTIQUE THÉORIQUE ET APPLIQUÉE, L'ÉCONOMÉTRIE

R. FERON

SUR LES TABLEAUX DE CORRÉLATION DONT LES MARGES SONT DONNÉES

R. FÉRON

SUR LES DISTRIBUTIONS DE PROBABILITÉ DE LAURENT SCHWARTZ

M. GIRAULT

PRODUIT DE FONCTIONS CARACTÉRISTIQUES

S. K. NASR

DES POSITIONS TYPIQUES D'UNE VARIABLE ALÉATOIRE

A. WINTNER

DES DISTRIBUTIONS SYMÉTRIQUES A FONCTIONS CARACTÉRISTIQUES CONVEXES

ANALYSE D'ARTICLES

VOL. V - FASCICULE 1 - 1956

PARIS

11, Rue Pierre Curie



SUR LES TABLEAUX DE CORRÉLATION DONT LES MARGES SONT DONNÉES

CAS DE L'ESPACE A TROIS DIMENSIONS[1]

par

R. FÉRON

Considérons l'ensemble E des fonctions de répartition à n dimensions F(x) et convenons d'écrire :

$$F(x) \prec G(x) \qquad \text{(relation } \mathcal{O}\text{)} \qquad (1)$$

si pour tout vecteur $\vec{x} \in R^n$

$$F(\vec{x}) \leqslant G(\vec{x}) \qquad (2)$$

Il est connu que la relation (1) est une relation d'ordre partiel car elle est :

Réflexive : on a pour tout $\vec{x}$ $\qquad F(\vec{x}) \leqslant F(\vec{x})$

Antisymétrique : Si on a pour tout $\vec{x}$:

$$F(\vec{x}) \leqslant G(\vec{x}) \quad \text{et} \quad F(\vec{x}) \geqslant G(\vec{x})$$

alors : $\qquad F(\vec{x}) \equiv G(\vec{x})$

Transitive : Si on a pour tout $\vec{x}$

$$F(\vec{x}) \leqslant G(\vec{x}) \quad \text{et} \quad G(\vec{x}) \leqslant (H(\vec{x}) \quad \text{alors} \quad F(\vec{x}) \leqslant H(\vec{x})$$

Ceci posé Birkhov a montré que :

Théorème de Birkhov

Si x est à une seule dimension, alors l'ensemble E des fonctions de répartition constitue un treillis :

En effet, quels que soient F(x) et G(x) ces éléments ont une borne supérieure qui n'est autre que

$$\text{Max} \left\{ F(x) \;,\; G(x) \right\} = F^{(1)} (x)$$

(1) Nous tenons ici à exprimer notre reconnaissance à M. J. Bass qui nous a donné l'idée d'étudier le cas tridimensionnel.

et une borne inférieure qui n'est autre que :

$$\text{Min}\left\{ F(x) \ , \ G(x) \right\} = F^{(2)} (x)$$

$F^{(1)}$ est bien une fonction de répartition car :

$$F^{(1)} (-\infty) = 0 \qquad\qquad F^{(1)} (+\infty) = 1$$

$$F^{(1)} (x + \Delta x) = \text{Max} \left[F(x+\Delta x), G(x + \Delta x) \right] \geqslant \text{Max} \left[F(x), G(x) \right] = F^{(1)} (x)$$

Par contre dans le cas où x est à 2 dimensions ou plus l'ensemble E n'est plus un treillis ni même un demi treillis, car :

1°) Il existe des fonctions de répartition égales à :

$$\text{Max}\left\{ F(x,y) \ , \ G(x,y)\right\} \quad \text{en} \quad x_0\, y_0$$

et supérieures (ou égales) à cette quantité en tout autre point, par exemple :

$F^{(2)} (x,y) =$

a) 1 pour $x > x_0$ et $y > y_0$

b) $\text{Max}\left\{F(x_0,y_0), G(x_0,y_0)\right\}$ pour $x_0 > x > x_1$ et $y > y_1$ ou $x > x_1$ et $y_0 > y > y_1$

où $x_1 y_1$ est un nombre tel que l'on ait simultanément :

$F(x_1,\infty)$, $G(x_1,\infty)$, $F(\infty,y_1)$, $G(\infty,y_1)$

inférieurs à $\text{Max}\left[F(x_0,y_0) \ , \ G(x_0,y_0)\right]$

c) $\text{Min}\left\{ \text{Max}\left[F(x,\infty) \ , \ G(x,\infty)\right], \ \text{Max}\left[(F(\infty,y) \ , \ G(\infty,y)\right]\right.$

pour $x < x_1$ ou $y < y_1$

Donc s'il existe une borne supérieure, ce sera nécessairement la fonction :

$$H(x,y) = \text{Max}\left\{ F(x,y) \ , \ G(x,y)\right\}$$

2°) Montrons que H n'est pas nécessairement une fonction de répartition. En effet pour qu'il en soit ainsi, il faudrait que pour tout couple F et G, la quantité

$$= H(x+\Delta x,y+\Delta y) - H(x+\Delta x,y) - H(x,y+\Delta y) + H(x,y)$$

soit nécessairement positive ou nulle.

Or, si :

$$F(x+\Delta x,y+\Delta y) = F(x+\Delta x,y) > G(x+\Delta x,y+\Delta y)$$

bien que :

$$G(x,y+\Delta y) > F(x,y+\Delta y) \qquad \text{et} \qquad G(x,y+\Delta y) > G(x,y)$$

alors on a

$$J^{(2)} = -G(x,y+\Delta y) + \text{Max}\left[F(x,y) \ , \ G(x,y)\right] \geqslant 0 \tag{3}$$

Donc H n'est pas une fonction de répartition et par conséquent, il n'existe pas de borne supérieure. Deux éléments quelconques n'ayant pas

nécessairement de borne supérieure, l'ensemble E n'est pas un treillis. On démontrerait de même que 2 éléments quelconques de E n'ont pas nécessairement de borne inférieure et par conséquent que E n'est pas nécessairement un demi-treillis.

Or, le théorème de Birkhov nous est apparu comme lié au théorème suivant dû à Fréchet [1].

Théorème de Fréchet.- L'ensemble E (ordonné par $\mathcal{O}$) des fonctions de répartition $F(x,y)$ de R_2 ayant des marges données : $A(x) = F(x,\infty)$, $B(y) = F(\infty,y)$, a un élément maximum $F^{(1)}(x,y)$ et un élément minimum $F^{(2)}(x,y)$.

Or, nous venons de montrer que si $F(x,y) \in E$ l'ensemble E ordonné $\mathcal{O}$ n'est pas un treillis. Il y a donc lieu de penser que le théorème de Fréchet ne se généralise pas aux espaces Euclidiens à plus de deux dimensions.

Examinons donc ce qui se passe dans l'espace à 3 dimensions. Ici 2 cas sont à considérer :

TABLEAUX DE CORRÉLATION OU LES MARGES ($F(x,\infty,\infty)$, $F(\infty,y,\infty)$ $F(\infty,\infty,z)$) SONT DONNEES

a) méthode brutale

Considérons maintenant l'ensemble E'_3 des fonctions de répartition $F(x,y,z)$ ayant des marges données $A(x) = F(x,\infty,\infty)$, $B(y) = F(\infty,y,\infty)$, $C(z) = F(\infty,\infty,z)$ et supposons E'_3 ordonné.

Montrons que cet ensemble a un élément maximum et en général pas d'élément minimum.

1. IL Y A UN ÉLÉMENT MAXIMUM

On aura évidemment

$$F(x, y, z) \leq \phi(x,y,z)$$

en posant :

$$\phi(x,y,z) = \text{Min}\left[A(x),\ B(y),\ C(z)\right]$$

Il est clair que si nous parvenons à montrer que ϕ est une fonction de répartition, ce sera un élément maximum. Or, il en est bien ainsi, en effet :

1°) $$\phi(-\infty,y,z) \leq A(-\infty)$$

donc :

$$0 \leq \phi(-\infty,y,z) \leq 0 \qquad \text{et} \qquad \phi(-\infty,y,z) = 0$$

(1) Fréchet. Sur les tableaux de corrélation dont les marges sont données. Ann. Univ. Lyon, 9, 1951, p. 53-77.

Il en est de même de

2°) $$\phi(+\infty,+\infty,+\infty) = \text{Min}\left\{ A(+\infty),\ B(+\infty),\ C(+\infty)\right\} = 1$$

3°) Quels que soient Δx, Δy, Δz positifs ou nuls, on a :

$$\begin{aligned} I^{(3)} = {} & \phi(x+\Delta x,y+\Delta y,z+\Delta z) - \phi(x,y+\Delta y,z+\Delta z) - \phi(x+\Delta x,y,z+\Delta z) \\ & - \phi(x+\Delta x,y+\Delta y,z) + \phi(x+\Delta x,y,z) + \phi(x,y,\Delta y,z) + \phi(x,y,z+\Delta z) - \phi(x,y,z) \\ = {} & I^{(2)}(x,\Delta x,y,\Delta y,z+\Delta z) - I^{(2)}(x,\Delta x,y,\Delta y,z) \end{aligned}$$

en posant

$$I^{(2)}(x,\Delta x,y,\Delta y,z) = \phi(x+\Delta x,y+\Delta y,z) - \phi(x+\Delta x,y,z) - \phi(x,y+\Delta y,z) + \phi(x,y,z)$$

De même, on a :

$$I^{(2)}(x,\Delta x,y,\Delta y,z) = I^{(1)}(x,\Delta x,y+\Delta y,z) - I^{(1)}(x,\Delta x,y,z)$$

en posant :

$$I^{(1)}(x,\Delta x,y,z) = \phi(x+\Delta x,y,z) - \phi(x,y,z)$$

Sans restreintre la généralité on peut supposer qu'au point x,y,z on a :

$$A(x) \leqslant B(y) \leqslant C(z)$$

On aura alors

$$I^{(1)}(x,\Delta x,y,z) = \text{Min}\left[A(x+\Delta x),\ B(y)\right] - A(x)$$

$$\begin{aligned} I^{(2)}(x,\Delta x,y,\Delta y,z) &= I^{(1)}(x,\Delta x,y+\Delta y,z) - I^{(1)}(x,\Delta x,y,z) \\ &= \text{Min}\left[A(x+\Delta x),B(y+\Delta y),C(z)\right] - A(x) - \left\{\text{Min}\left[A(x+\Delta x),B(y)\right] - A(x)\right\} \\ &= \text{Min}\left[A(x+\Delta x),B(y+\Delta y),C(z)\right] - \text{Min}\left[A(x+\Delta x),B(y)\right] \end{aligned}$$

$$\begin{aligned} I^{(3)} &= I^{(2)}(x,\Delta x,y,\Delta y,z+\Delta z) - I^{(2)}(x,\Delta x,y,\Delta y,z) \\ &= \text{Min}\left[A(x+\Delta x),B(y+\Delta y),C(z+\Delta z)\right] - \text{Min}\left[A(x+\Delta x),B(y+\Delta y),C(z)\right] \geqslant 0 \quad (4) \end{aligned}$$

quantité évidemment positive ou nulle. Donc $\phi(x,y,z)$ est bien une fonction de répartition et est par conséquent un élément maximum.

2. IL N'Y A PAS EN GÉNÉRAL D'ÉLÉMENT MINIMUM

En effet, on a :

$$1 - F(x,y,z) \leqslant \text{Min}\left[(1-A(x) + 1 - B(y) + 1 - C(z),\ 1\right] \quad (5)$$

et on peut avoir l'égalité pour $F(x,y,z)$ convenablement choisi pour un $x_0\, y_0\, z_0$ donné mais à part cela arbitraire. (On voit en effet que l'égalité

aura lieu si on a simultanément $Pr(X>x_0, Y>y_0) = 0$, $Pr(Y>y_0, Z>z_0) = 0$, $Pr(Z>z_0, X>x_0) = 0$

Or, de (5), on tire :

$$F(x,y,z) \geqslant \text{Max}\left[A(x) + B(y) + C(z) - 2,\ 0\right] \qquad (6)$$

cette borne inférieure pouvant effectivement être atteinte pour certains éléments $F(x,y,z)$. Donc s'il existe un élément minimum ce ne pourra être que :

$$\psi(x,y,z) + \text{Max}\left\{A(x) + B(y) + C(z) - 2,\ 0\right\} \qquad (7)$$

Nous allons montrer que $\psi(x,y,z)$ n'est pas nécessairement une fonction de répartition parce que la qualité $I^{(3)}$ relative à ψ n'est pas nécessairement positive.

En effet, nous avons

$$\begin{aligned} I^{(3)} = {} & \psi(x+\Delta x, y+\Delta y, z+\Delta z) - \psi(x+\Delta x, y+\Delta y, z) - \psi(x, y+\Delta y, z+\Delta z) \\ & - \psi(x+\Delta x, y, z+\Delta z) \\ & + \psi(x+\Delta x, y, z) + \psi(x, y+\Delta y, z) + \psi(x, y, z+\Delta z) - \psi(x,y,z) \end{aligned}$$

Supposons que x, y, z, Δx, Δy, Δz aient été choisis tels que :

$$\psi(x+\Delta x, y, z) = \psi(x, y+\Delta y, z) = \psi(x, y, z+\Delta z) = 0 \qquad (8)$$

alors que :

$$\psi(x+\Delta x, y+\Delta y, z) > 0 \qquad \psi(x, y+\Delta y, z+\Delta z) > 0 \qquad \psi(x+\Delta x, y, z+\Delta z) > 0$$

on aura alors

$$\begin{aligned} I^{(3)} = {} & A(x+\Delta x)+B(y+\Delta y)+C(z+\Delta z)-2-\left\{A(x+\Delta x)+B(y+\Delta y)+C(z)-2+A(x+\Delta x)\right. \\ & \left. + B(y)+C(z+\Delta z)-2+A(x)+B(y+\Delta y)+C(z+\Delta z)-2\right\} \\ = {} & -\left[A(x+\Delta x)+B(y+\Delta y)+C(z)-2\right]-\left[A(x)+B(y)+C(z+\Delta z)-2\right] \qquad (9) \end{aligned}$$

La première parenthèse est positive d'après les hypothèses faites. Si donc on a choisi Δz de manière que la seconde soit nulle, ce qui est compatible avec (8) et toujours possible dans le cas continu, alors $I^{(3)}$ est négative, ce qui montre que ψ n'est pas une fonction de répartition en général.

Autre méthode : valable seulement dans le cas continu

Sur $F(x,y,z)$ faisons le changement de variables $\xi = A(x)$, $\eta = B(y)$, $\zeta = C(z)$. Il est connu que dans ces conditions la fonction de répartition $F(x,y,z)$ se transforme en une fonction de répartition $G(\xi,\eta,\zeta)$ telle que $G(\xi,\infty,\infty) = \xi$ pour $0 \leqslant \xi \leqslant 1$, $G(\infty,\eta,\infty) = \eta$ $(0 \leqslant \eta \leqslant 1)$, $G(\infty,\infty,\zeta) = \zeta$ $(0 \leqslant \zeta \leqslant 1)$. Étendant légèrement une dénomination de Georges Darmois, on peut appeler $G(\xi,\eta,\zeta)$ la carte de $F(x,y,z)$ sur le cube unité.

De plus, on voit immédiatement que la transformation considérée conserve l'ordre $\mathcal{O}$ défini par (2). Autrement dit, étant données 2 fonctions de répartition $F^{(1)}(x,y,z)$ et $F^{(2)}(x,y,z)$ ayant des marges données $A(x)$, $B(y)$, $C(z)$, la relation $F^{(1)}(x,y,z) \leqslant F^{(2)}(x,y,z)$ pour tout point x,y,z entraine la relation similaire sur les transformées $G^{(1)}(\xi,\eta,\zeta)$ et $G^{(2)}(\xi,\eta,\zeta)$ de $F^{(1)}$ et $F^{(2)}$, soit $G^{(1)}(\xi,\eta,\zeta) \leqslant G^{(2)}(\xi,\eta,\zeta)$ pour tout point ξ,η,ζ. La réciproque est d'ailleurs exacte. On a donc en définitive :

$$F^{(1)}(x,y,z) < F^{(2)}(x,y,z) \rightleftarrows G^{(1)}(\xi,\eta,\zeta) < G^{(2)}(\xi,\eta,\zeta) \qquad (10)$$

La recherche des éléments maximaux et minimaux sur les fonctions de répartition $F(x,y,z)$ ayant des marges données est donc complètement équivalente à la recherche des mêmes éléments maximaux et minimaux lorsque les variables marginales sont réparties uniformément sur les segments 0, 1 des divers axes. Donc, si nous arrivons à montrer que les fonctions G admettent un élément maximum et plusieurs éléments minimaux, il en sera nécessairement de même des fonctions $F(x,y,z)$ ayant des marges données.

1°- LES $G(\xi,\eta,\zeta)$ ONT UN ÉLÉMENT MAXIMUM

En effet, on a évidemment : $G(\xi,\eta,\zeta) \leqslant \operatorname{Min}[\xi,\eta,\zeta]$ et $H(\xi,\eta,\zeta) = \operatorname{Min}[\xi,\eta,\zeta]$ est bien une fonction de répartition car :

1) $H(-\infty,\eta,\zeta) = H(\xi,-\infty,\zeta) = H(-\infty,\eta,\zeta) = 0$

2) $H(+\infty,+\infty,+\infty) = 1$

3) $$\begin{aligned} I^{(3)} = {} & H(\xi+\Delta\xi,\eta+\Delta\eta,\zeta+\Delta\zeta) - H(\xi+\Delta\xi,\eta+\Delta\eta,\zeta) \\ & - H(\xi,\eta+\Delta\eta,\zeta+\Delta\zeta) + H(\xi+\Delta\xi,\eta,\zeta) + H(\xi,\eta+\Delta\eta,\zeta) \\ & + H(\xi,\eta,\zeta+\Delta\zeta) - H(\xi,\eta,\zeta) \end{aligned}$$

est facile à évaluer si nous supposons par exemple $\xi < \eta < \zeta$. On peut alors répéter mot pour mot la démonstration de la page 4 avec les nouvelles notations (c'est-à-dire en remplaçant $A(x)$ par ξ, $B(y)$ par η, $C(z)$ par ζ et on a $I^{(3)} \geqslant 0$.

Support de l'élément maximum

Il est intéressant de caractériser cette distribution au moyen de la densité généralisée $\frac{\partial^3 H}{\partial x \partial y \partial z}$ qui se définit exactement comme pour 2 dimensions (1)

On a :

$$< \frac{\partial^3 H}{\partial x \partial y \partial z}, \varphi > = - < H, \frac{\partial^3 \varphi}{\partial x \partial y \partial z} > = \iiint \varphi \, dH(x,y,z) \qquad (11)$$

(1) Voir Féron. Sur les distributions de probabilité de Laurent Schwartz et certaines de leurs applications au calcul des probabilités.

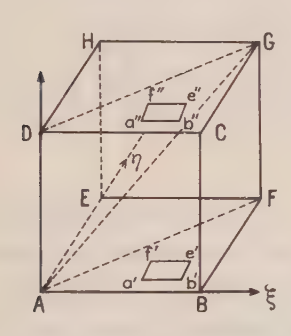
Montrons que :

Théorème = Le support de $\frac{\partial^3 H}{\partial x \partial y \partial z}$ est la droite AG.

Nous donnerons 3 démonstrations différentes de ce théorème.

1/ Démonstration géométrique

Il suffit, somme toute de montrer que sur tout parallélépipède ne rencontrant pas AG, $I^{(3)} = \iiint d\,H(x,y,z) = 0$. Donc en définitive que le $I^{(3)}$ relatif à H est nul sur tout tel parallélépipède a b c d e f g h qui ne coupe pas AG. Ceci peut se traduire en disant qu'au moins une des projections du parallélépipède sur les plans $\xi A \eta$, $\xi A \zeta$, $\eta A \zeta$ ne rencontre pas la première bissectrice correspondante. Supposons par exemple que a b e f se projette en a'b'e'f' et que a'b'e'f' ne rencontre pas AF. Alors si a''b''e''f'' sont les projections des mêmes points sur le plan CDH, il est clair que l'intégrale I étendue à a b c d e f g h est inférieure à la même intégrale étendue à a'b'e'f'a''b''e''f''. On est donc ramené à un problème à deux dimensions dans le plan des $\xi\,\eta$. La fonction de répartition étant Min $(\xi,\eta) = H^{(1)}(\xi,\eta)$ quelle est la probabilité pour que le point M se trouve dans le rectangle a'b'e'f' ? Cette probabilité étant nulle, comme nous l'avons montré dans l'article cité précédemment, il en résulte que $I^{(3)} = 0$, donc que AG est le support de $\frac{\partial^3 H}{\partial x \partial y \partial z}$



2/ Démonstration algébrique

Il suffit somme toute de traduire en langage algébrique la démonstration précédente. Considérons le parallélépipède a b c d e f g h comme défini par les 2 points $a(\xi,\eta,\zeta)$ et $g(\xi+\Delta\xi, \eta+\Delta\eta, \zeta+\Delta\zeta)$. Le fait que a b c d e f g h ne coupe pas AG se traduit par le fait qu'on ne peut trouver λ tel que $0 < \lambda < 1$ et que $\xi + \lambda\Delta\xi = \eta + \lambda\Delta\eta = \zeta + \lambda\Delta\zeta$. Supposons par exemple que c'est la première de ces inégalités $\xi + \lambda\Delta\xi = \eta + \lambda\Delta\eta$ qui ne peut être vérifiée, $\frac{\partial^3 H}{\partial x \partial y \partial z}$ étant positive, on aura :

$$0 \leqslant I^{(3)} \leqslant \int_{\xi}^{\xi+\Delta\xi}\int_{\eta}^{\eta+\Delta\eta}\int_{0}^{1} dH(x,y,z) = I'^{(3)}$$

Or :

$$I'^{(3)} = \text{Min}\,(\xi+\Delta\xi, \eta+\Delta\eta, 1) - \text{Min}\,(\xi, \eta+\Delta\eta, 1) - \text{Min}\,(\xi+\ \xi, \eta, 1) + \text{Min}\,(\xi, \eta, 1)$$

$$- \{ \text{Min}\,(\xi+\Delta\xi, \eta+\Delta\eta, 0) - \text{Min}\,(\xi, \eta+\Delta\eta, 0) - \text{Min}\,(\xi+\Delta\xi, \eta, 0) + \text{Min}\,(\xi, \eta, 0)$$

$$= \text{Min}\,(\xi+\Delta\xi, \eta+\Delta\eta) - \text{Min}\,(\xi, \eta+\Delta\eta) - \text{Min}\,(\xi+\Delta\xi, \eta) + \text{Min}\,(\xi, \eta)$$

$$= 0$$

3/ <u>Autre démonstration</u>

On a :

$$< \frac{\partial^3 G(x,y,z)}{\partial x\, \partial y\, \partial z}, \; 1_{0,1}^{(z)} > = \frac{\partial G(x,y,\infty)}{\partial x\, \partial y} \qquad (12)$$

en appelant $1_{0,1}^{(z)}$ la fonction φ (z) égale à 1 sur l'intervalle 0 1 à 0 ailleurs. (fonction caractéristique de l'intervalle 0 1).

Il en résulte que le support $\frac{\partial^3 G(x,y,z)}{\partial x\, \partial y\, \partial z}$ devra nécessairement être contenu dans le produit de l'ensemble support de $\frac{\partial^2 G(x,y,\infty)}{\partial x\, \partial y}$ et de l'intervalle 0,1 porté par Oz. Dans le cas particulier de H(x,y,z), le support de $\frac{\partial^2 H(x,y,\infty)}{\partial x\, \partial y}$ est le segment de droite DG et l'ensemble produit de ce support par le segment DA est le rectangle A F G D. On démontrerait de même que le support de $\frac{\partial^3 H}{\partial x\, \partial y\, \partial z}$ doit être contenu dans le rectangle ACGE et dans le rectangle A B G H. Ce support devra donc être tout entier contenu sur le segment de droite AG. Pour montrer que ce support est le segment AG lui-même, il suffit d'appliquer l'opérateur $\frac{\partial^3 H}{\partial x\, \partial y\, \partial z}$ à une fonction φ (x,y,z) égale à 1 si $0 < \xi < x < \xi + \Delta\xi < 1$ à 0 ailleurs et de montrer que $< \frac{\partial^3 H}{\partial x\, \partial y\, \partial z}, \varphi >$ est différent de zéro et ceci quels que soient ξ et $\Delta\xi$. En effet, pour une telle fonction φ, on a toujours quel que soit G, telle que G $(\xi, \infty, \infty) = \xi$

$$< \frac{\partial^3 G}{\partial x\, \partial y\, \partial z}, \varphi > = \int_{\xi}^{\xi+\Delta\xi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} dG = \Delta\xi \qquad (13)$$

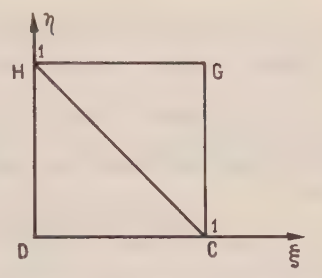
2°/ IL N'Y A PAS D'ÉLÉMENT MINIMUM

<u>1ère démonstration</u>

En effet, supposons qu'il existe un élément minimum $K(\xi, \eta, \zeta) \in E'_3$, c'est-à-dire un élément tel que $K(\xi, \eta, \zeta) \leq G(\xi, \eta, \zeta)$ pour tout $G \in E'_3$ et quels que soient ξ, η, ζ ; en particulier, on devra avoir :

$$K(\xi, \eta, \infty) \leq G(\xi, \eta, \infty)$$

pour tout $G \in E'_3$ et quels que soient ξ et η. C'est dire que dans l'espace 2 dimensions $K(\xi, \eta, \infty)$ devra être l'élément minimum des fonctions de répartition $\varphi(\xi, \eta)$ telles que $\varphi(\xi, \infty) = \xi$ $\varphi(\infty, \eta) = \eta$. Cet élément est bien connu. On a $K(\xi, \eta, \infty) = \mathrm{Max}(0, \xi+\eta-1)$ et nous avons montré dans l'article cité que le support de la distribution $\frac{\partial^2 K(\xi, \eta, \infty)}{\partial \xi \partial \eta}$ est le segment de droite HC. Et comme précédemment on voit d'après (12) que le support de K devrait être contenu dans le plan HCEB.



En raisonnant de même sur $K(\infty, \eta, \zeta)$ et $K(\xi, \infty, \zeta)$ on verrait que le support doit être tout entier contenu dans les rectangles DEFC et DBFH. Il en résulte que le support de K devrait nécessairement se trouver à l'intersection de ces 3 plans, c'est-à-dire se réduire au seul point O centre du cube. Mais d'après (13) le support de tout G doit nécessairement couper le plan $\xi = \xi_0$ pour tout ξ_0 $(0 < \xi_0 < 1)$. On aboutit donc à une impossibilité. Ce qui prouve que notre hypothèse initiale est fausse. Autrement dit, il n'existe pas d'élément minimum.

Autre démonstration

Pour montrer qu'il n'existe pas d'élément minimum, il suffit de montrer qu'il existe au moins 2 éléments minimaux distincts. Or, il est clair que si nous avons pour un certain $K^{(1)}(\xi, \eta, \zeta)$, $K^{(1)}(\xi, \eta, \infty) = \mathrm{Max}(0, \xi+\eta-1)$, on devra avoir la même relation pour tout $\varphi(\xi, \eta, \zeta) < K^{(1)}(\xi, \eta, \zeta)$. Il en résulte que le support de tout élément inférieur à $K^{(1)}$ doit être contenu dans le plan C H E B. Si nous imposons de plus à $K^{(1)}(\infty, \eta, \zeta)$ d'être minimum, c'est-à-dire que $K^{(1)}(\infty, \eta, \zeta) = \mathrm{Max}(0, \eta+\zeta-1)$, alors le support de tout élément φ inférieur à $K^{(1)}$ devra être tout entier contenu sur l'intersection des plans DECF et CHEB, c'est-à-dire sur la droite EC. Mais en vertu de 13, il n'y a qu'une seule distribution $K^{(1)}$ ayant pour support la droite EC. C'est la distribution telle que $P(S) = \frac{\text{longueur de la portion de EC} \subset S}{\text{longueur du segment EC}}$ donc il n'existe pas de fonction de répartition φ inférieure à $K^{(1)}$ et par conséquent $K^{(1)}$ est bien une fonction de répartition minimale. On démontrerait de même que les fonctions de répartition $K^{(2)}$ et $K^{(3)}$ ayant pour support les droites DF et HB sont aussi minimales. Il en résulte directement qu'il n'existe pas d'élément minimum.

Remarque : Comme dans le cas bidimensionnel, toute fonction de répartition $\varphi(\xi, \eta, \zeta)$ qui vérifie en tous points les inégalités :

$$\mathrm{Max}\left[0, \xi+\eta+\zeta-2\right] \leq \varphi(\xi, \eta, \zeta) \leq \mathrm{Min}(\xi, \eta, \zeta)$$

a les marges voulues. La seule différence avec le cas bidimensionnel est que dans l'espace à 4 dimensions la surface séparatrice inférieure n'est plus une fonction de répartition.

II. CAS OU ON CONNAIT $F(x, y, \infty)$ et $F(\infty, \infty, z)$

Pour montrer qu'il n'y a en général ni élément minimum ni élément maximum, il nous suffira de donner deux contre-exemples sur le cube.

a) Il n'y a pas d'élément minimum en général

En effet, si nous prenons pour $F(x,y,\infty)$ la fonction

$$G(\xi,\eta) = \text{Max}\,(0,\xi+\eta-1) \text{ et } F(\infty,\infty,\zeta) = \zeta \qquad \text{pour } 0 \leqslant \zeta \leqslant 1$$

alors il existe 2 fonctions de répartition minimales dont les supports sont HB et EC.

b) Il n'y a pas d'élément maximum en général

En effet, avec l'exemple précédent, les 2 fonctions de répartition considérées sont aussi maximales.

SUR LES DISTRIBUTIONS DE PROBABILITÉ DE LAURENT SCHWARTZ

ET QUELQUES UNES DE LEURS APPLICATIONS AU CALCUL DES PROBABILITÉS

par

R. FÉRON

SOMMAIRE : La distribution de probabilité de Laurent Schwartz est une sorte de densité de probabilité généralisée.

On démontre en effet que la dérivée d'une fonction de répartition (prise au sens des distributions) est une distribution de probabilité au sens de Laurent Schwartz et que réciproquement toute distribution de probabilité est la dérivée d'une certaine fonction de répartition. De sorte que la définition de la probabilité de Schwartz est complètement identique à la définition classique de la probabilité comme fonction complètement additive d'ensembles.

Nous verrons toutefois dans le cours de ce travail qu'il est plus simple et plus intuitif pour la démonstration de certains théorèmes d'adopter la définition de Schwartz plutôt que la définition classique. On voit en effet mieux ce qui se passe en raisonnant sur les distributions plutôt que sur les fonctions de répartition.

Ce théorème permet notamment d'étudier les couples aléatoires ayant des distributions marginales données. Certains d'entre eux sont définis par des distributions ayant leur support tout entier contenu sur des courbes que l'on définit d'une manière précise.

PLAN DU TRAVAIL

1°) Définitions fondamentales

2°) Exemples

3°) Généralisation à l'espace à n dimensions

4°) Dérivées d'une Distribution

5°) Application à l'étude des couples aléatoires dont les distributions marginales sont connues

6°) Application au problème de Bertaut.

1. - DÉFINITIONS FONDAMENTALES

Laurent Schwartz (1) a caractérisé toute variable aléatoire X au moyen d'une certaine distribution particulière notée T ou T_x et appelée par lui sa distribution de probabilité. Aussi après avoir rappelé brièvement la définition des distributions de Laurent Schwartz indiquerons-nous en détail les conditions que doit vérifier une distribution pour lui conférer la qualité de distribution de probabilité.

I. Définition des distributions de Laurent Schwartz

Définition : On appelle distribution une fonctionnelle linéaire continue sur l'espace vectoriel $\mathcal{D}$ des fonctions indéfiniment dérivables à support borné.

Cela signifie qu'à toute fonction $\varphi(x)$ indéfiniment dérivable (2) et nulle en dehors d'un intervalle borné (si $x \in R^1$) ou d'un pavé borné (si $x \in R^n$) on associe un nombre noté $T(\varphi)$ ou $< T, \varphi >$ jouissant des propriétés suivantes :

1°) T est une fonctionnelle linéaire, c'est-à-dire que :

$$< T, \varphi_1 + \varphi_2 > = < T, \varphi_1 > + < T, \varphi_2 >$$

et :

$$< T, \lambda \varphi > = \lambda < T, \varphi >$$

2°) T est continu sur $\mathcal{D}$, autrement dit $< T, \varphi_j > \rightarrow < T, \varphi >$ (3) si la suite de fonctions φ_j converge vers φ au sens de la pseudo-topologie de $\mathcal{D}$, c'est-à-dire si :

a) les φ_j sont nuls en dehors d'un intervalle (ou d'un pavé) borné indépendant de j.

b) les dérivées de tout ordre m des φ_j convergent uniformément vers les dérivées correspondantes de φ (4)

---

(1) cf. Laurent Schwartz cours de MMP 1954 fasc. 4, p. 15.

(2) Il existe effectivement de telles fonctions, par exemple $\varphi(x) = 0$ si $|x| \geqslant 1$

$\varphi(x) = e^{-\frac{1}{1-x^2}}$ si $|x| < 1$

(3) Cette dernière définition est évidemment complètement équivalente à celle donnée par Schwartz dans son livre : Théorie des distributions (p. 21 et 24). Il dit en effet ceci :

1°) Nous appellerons $\mathcal{D}$ l'espace vectoriel des fonctions complexes $\varphi(x) = \varphi(x_1, \ldots x_n)$ de n variables réelles indéfiniment dérivables à support compact...

2°) Une distribution T sera alors une forme linéaire continue sur $\mathcal{D}$. $T(\varphi)$ est continue signifie que : Si des $\varphi_j \in \mathcal{D}$ ont leurs supports contenus dans un compact fixe de R^n et si elles convergent uniformément vers 0 dans R^n ainsi que chacune de leurs dérivées, alors les nombres complexes $T(\varphi_j)$ convergent vers zéro.

(4) On remarquera que Laurent Schwartz a imposé aux fonctions φ et à la pseudotopologie sur $\mathcal{D}$ les conditions les plus restrictives possibles. L'intérêt de ces restrictions tient dans le fait que $\mathcal{D}$ étant le plus petit possible son dual sera le plus grand possible et les théorème obtenus sur les distributions sont les plus généraux possibles. Il existe évidemment d'autres théories obtenues en remplaçant l'espace $\mathcal{D}$ par des espaces moins généraux. Telle est par exemple la théorie de la mesure où l'espace $\mathcal{D}$ est remplacé par l'espace $\mathcal{C}$ des fonctions continues. Les théorèmes obtenus à l'aide de ces théories seront parfois plus puissants que ceux obtenus dans la théorie des distributions. Ils seront toujours moins généraux.

II. Distributions de probabilité de Laurent Schwartz

Laurent Schwartz attache à toute variable aléatoire (v.a) une certaine distribution T_X appelée sa distribution de probabilité (1) et qui devra être positive et telle que $< T, 1_{-\infty+\infty} > = 1$.

Cette définition comporte les trois points suivants :

a/ T est une distribution (voir définition plus haut)

b/ La distribution T est positive.

C'est-à-dire que pour toute fonction $\varphi(x) \geqslant 0$ pour tout x le nombre $< T, \varphi >$ est un nombre positif.

c/ $< T, 1_{-\infty,+\infty} > = 1$

Cette définition peut surprendre, attendu que $1_{-\infty,+\infty}$ fonction identique à 1 sur tout l'intervalle $-\infty, +\infty$ n'est pas à support borné.

Laurent Schwartz arrive pourvant à sa définition en se basant sur les théorème suivants :

Théorème I. - Toute fonction continue $\chi(x)$ à support (2) borné K, peut être uniformément approchée à ε près quel que soit $\varepsilon > 0$ par une fonction indéfiniment dérivable $\varphi \in \mathcal{D}$ et on peut astreindre $\varphi(x)$ à avoir son support contenu dans un voisinage arbitraire du support K de $\chi(x)$.

Théorème II. - (Fréchet (3)). - Si $\psi(x)$ est une fonction mesurable, il existe une suite de fonctions $\chi_j(x)$ continues qui pour j tendant vers l'infini convergent presque partout vers $\psi(x)$.

Théorème III. - Si nous convenons de désigner par $< f, \varphi >$ et $< g, \varphi >$ les fonctionnelles $\int_{-\infty}^{+\infty} f(x)\,\varphi(x)\,dx$, $\int_{-\infty}^{+\infty} g(x)\,\varphi(x)\,dx$ (f et g localement sommables (4) et à part cela quelconques).

Alors on ne peut avoir $< f, \varphi > = < g, \varphi >$ (1) pour tout $\varphi \in \mathcal{D}$ que si $f = g$ presque partout.

Pour prouver ce théorème Schwartz montre que :

a) si (1) a lieu pour une suite de φ_j convergeant vers une fonction continue à support borné χ, alors d'après le théorème I :

$$\int f(x)\,\chi(x)\,dx = \int g(x)\,\chi(x)\,dx \qquad (2)$$

---

(1) On notera que la définition de la distribution de probabilité de Laurent Schwartz ne coincide pas avec celle de la probability distribution des Anglo-Saxons. On peut toutefois montrer que la connaissance de la distribution de probabilité de Schwartz entraine celle de la probability distribution et réciproquement, de sorte que le terme ne risque pas d'induire le lecteur en erreur.

(2) Rappelons qu'on appelle support d'une fonction continue f(x) l'adhérence de l'ensemble des points x(de R^1 ou de R^n) tels que $f(x) \neq 0$.

(3) Le théorème II démontré par Lebesgue dans le cas des fonctions sommables a été démontré pour le cas des fonctions mesurables par Maurice Fréchet dans sa thèse.

(4) C'est-à-dire sommables sur tout intervalle fini.

b) Si (2) a lieu pour une suite χ_j de valeur de χ qui pour j tendant vers l'infini convergent presque partout vers une certaine fonction mesurable ψ alors :

$$\int f(x)\, \psi(x)\, dx = \int g(x)\, \psi(x)\, dx \tag{3}$$

Ces théorèmes conduisent Schwartz à définir le "prolongement analytique" de $< T, \varphi >$ dans le champ plus étendu des fonctions continues à support borné, puis dans le champ des fonctions mesurables à support borné et enfin dans le champ des fonctions mesurables à support quelconque et à adopter les définitions suivantes :

a) On définit la quantité $< T, \chi >$ où χ est une fonction continue à support borné comme la limite si elle existe, d'une suite de nombres : $< T, \varphi_i >$ (cf. th. I).

b) On définit le nombre $< T, \psi >$ où ψ est une fonction mesurable à support borné comme la limite si elle existe d'une suite de nombres $< T, \chi_i >$

c) On définit $< T, \omega >$ où ω est une fonction mesurable quelconque comme la limite si elle existe d'une suite de nombres $< T, \psi_i >$

En particulier, le fait que $< T, 1_{-\infty,+\infty} > = 1$ signifie que si j'appelle $1_{x_1,x_2}$ la fonction identique à 1 sur l'intervalle x_1, x_2 nulle ailleurs (1), alors la limite de $< T, 1_{x_1,x_2} >$ quand $x_1 \to -\infty$ et $x_2 \to +\infty$ existe et que cette limite est précisément égale à 1.

Ceci étant, nous appellerons <u>probabilité</u> (2) pour que la v.a. X soit comprise entre deux nombres x_1 et x_2 la quantité :

$$\Pr(x_1 \leqslant X \leqslant x_2) = < T, 1_{x_1,x_2} > \tag{4}$$

2. - EXEMPLES

a) <u>Définition.- On dit qu'une distribution quelconque T peut être identifiée à la fonction f si quel que soit $\varphi \in \mathcal{D}$ on a</u> :

$$< T, \varphi > = \int f(x)\, \varphi(x)\, dx \tag{5}$$

Alors on a les 2 propositions suivantes :

1/ Si f est une densité de probabilité et si T peut être identifiée à f, alors T est une distribution de probabilité et (4) donne :

$$\Pr\left\{ x_1 < X < x_2 \right\} = \int_{x_1}^{x_2} f(x)\, dx \tag{5'}$$

2/ Réciproquement si T est une distribution de probabilité et peut être identifiée à une fonction f, alors f est nécessairement une densité de probabilité et on a (5')

b) Si on a $< T, \varphi > = \varphi(o)$ (6), alors $\Pr(X = 0) = 1$ nous dirons dans ce cas que la distribution T peut être identifiée avec la distribution de Dirac (3) δ.

(1) $1_{x_1,x_2}$ est souvent appelée l'indicatrice de l'intervalle x_1, x_2.

(2) On peut montrer qu'en vertu des hypothèses faites sur T la quantité définie par (4) est bien une probabilité au sens classique.

(3) La distribution de Dirac δ a été introduite par Laurent Schwartz pour éviter l'emploi de la fonction de Dirac $\delta(x)$ qui est définie comme possédant les 2 propriétés mathématiquement contradictoires d'être nulle partout sauf pour $x = 0$ et d'être telle que: $\int_{-\infty}^{+\infty} \delta(x)\, dx = +1$

c) Si on a $<T, \varphi> = \varphi(x^i)$ (7), alors $Pr(X = x^i) = 1$; nous dirons alors que T peut être identifiée à la "translatée" δ_{x^i} de la distribution de Dirac δ et nous écrirons $T = \delta_{x^i}$.

d) si on a $<T, \varphi> = \sum p_i \varphi(x_i)$ (8), alors $Pr(X = x^i) = p_i$ et on a $T = \sum p_i \delta_{x^i}$ (8)

Définition.- On dit que la distribution T quelconque, peut être identifiée à une mesure $\mu(x)$ si $<T, \varphi> = \int \varphi(x)\, d\mu(x)$ (9).

Alors on a les 2 propositions suivantes :

1/ Si $\mu(x)$ est une fonction de répartition et si T peut être identifiée à la mesure $\mu(x)$ alors T est nécessairement une distribution de probabilité et (4) donne :

$$Pr(x^1 < X \leq x^2) = \mu(x^2) - \mu(x^1) \qquad (9')$$

2/ Réciproquement si T est une distribution de probabilité et peut être identifiée à une mesure $\mu(x)$ alors $\mu(x)$ est nécessairement une fonction de répartition et on a (9').

3. - GÉNÉRALISATION A L'ESPACE A n DIMENSIONS

Nous pouvons généraliser à l'espace à n dimensions les résultats de Laurent Schwartz. A tout vecteur aléatoire $\vec{X}$ nous associons une distribution $T_{x_1 x_2 \cdots x_n}$ positive et telle que :

$$<T_{x_1 \quad x_n}, 1_{(-\infty, \dots -\infty), (+\infty, \dots +\infty)}> = 1 \qquad (10)$$

où $1_{(-\infty, \dots -\infty), (+\infty, \dots +\infty)}$ est la fonction identique à 1 sur tout l'espace.

Et nous appellerons probabilité pour que le vecteur aléatoire $\vec{X}$ soit dans le pavé $x(x^1, \dots x^n)$, $y(y^1, \dots y^n)$, le nombre $<T_{\xi_1 \cdots \xi_n}, 1_{(x^1, \dots x^n), (y^1, \dots y^n)}>$ où $1_{(x^1, \dots x^n), (y^1, \dots y^n)}$ est la fonction identique à 1 sur le pavé (x,y, à 0 ailleurs.

Avec cette définition de probabilité, la fonction de répartition s'écrira évidemment :

$$F(x) = <T_{\xi_1 \dots \xi_n}, 1_{(-\infty \dots -\infty), (x^1 \dots x^n)}> \qquad (11)$$

Ceci posé, Laurent Schwartz a montré (1) que :

Théorème IV.- Toute distribution positive peut être identifiée à une mesure positive (2).

(1) Laurent Schwartz théorie des distributions p. 28 th. V.

(2) La démonstration de ce théorème fondamental est basée sur le fameux théorème de Riesz (C.R. 148, 1909, P. 974-6) qui dit qu'à toute forme linéaire $<L, \varphi>$ continue sur l'espace $\mathcal{C}$ des fonctions continues, on peut associer une mesure bien déterminée et telle que $<\mu, \varphi> = <L, \varphi>$

Il en résulte que pour montrer le théorème IV, il suffit de montrer que si T est une distribution positive le fait que les φ_j convergent vers 0 dans $\mathcal{C}$ (c'est-à-dire sont tels que la suite des φ_j converge uniformément vers 0, mais pas la suite des $\varphi_j^{(n)}$) entraine à lui seul que la suite $< T, \varphi_j>$ converge vers 0.

Du théorème IV résulte immédiatement :

Théorème V.- Toute distribution de probabilités peut être identifiée à une fonction de répartition : $< T; \varphi > = \int \varphi(x) \, d\, F(x)$

Ayant donc montré l'intérêt de la définition de la probabilité de Schwarts, rappelons brièvement les quelques définitions et théorèmes fondamentaux qui nous sont indispensables pour travailler sur elle (1). Certains de ces théorèmes (ceux relatifs aux distributions en général) sont classiques. D'autres (ceux relatifs aux distributions de probabilité) sont nouveaux.

4. - DÉRIVÉES D'UNE DISTRIBUTION

a) Théorèmes sur les distributions en général.

1) Cas de l'espace à une dimension.

Définition : La dérivée de la distribution T est la fonctionnelle $S = \frac{dT}{dx}$ définie par :

$$< \frac{dT}{dx}, \varphi > = - < T, \frac{d\varphi}{dx} > \qquad (12)$$

$(\varphi \in \mathcal{D})$ (2).

Les théorèmes suivants sont presque évidents :

Théorème VI.- La dérivée d'une distribution est une distribution.

Théorème VII.- Si T peut être identifiée à la fonction f continue et dérivable en tous points, alors sa dérivée $\frac{dT}{dx}$ peut être identifiée à la fonction f', on écrit alors $\frac{df}{dx} = f'$ (13) et par abus de langage, on dit que si f est continue, sa dérivée au sens des distributions est égale à sa dérivée au sens des fonctions. Ainsi, si une variable X possède une densité de probabilité cette densité est la dérivée de la fonction de répartition F(x) de x au sens des distributions.

Par contre si la fonction F(x) présente des discontinuités sa dérivée $f(x) = F'(x)$ au sens des distributions n'est pas identique à $\{F'(x)\}$ distribution identifiée à la fonction (3) égale à la dérivée (au sens de la théorie des fonctions) de F(x) aux points de continuité.

---

(1) Nous ne rappellerons ici que les définitions et théorèmes dont nous avons effectivement besoin. D'autres définitions sont encore plus fondamentales en calcul des probabilités comme celles concernant le produit de convolution de deux distributions et la transformée de Fourier d'une distribution.

(2) L'expression (12) reste évidemment valable si $\varphi \in \mathcal{D}^{(1)}$ ensemble des fonctions dérivables, pour certains opérateurs T particuliers (notamment si T peut être identifié à une fonction.

(3) Dans cette théorie deux fonctions presque partout égales sont considérées comme identiques.

Exemple : Supposons que la v.a. X soit presque certainement nulle. Dans ces conditions, sa fonction de répartition est la fonction de Heaviside $\Upsilon(x)$ (lire Upsilon de x) telle que

$$\begin{aligned} \Upsilon(x) &= 0 \quad \text{si} \quad x < 0 \\ &= 1 \quad \text{si} \quad x \geq 0 \end{aligned} \qquad (14)$$

A cette fonction correspond la distribution de Heaviside que par convention nous désignons aussi par $\Upsilon(x)$ et qui est telle que :

$$< \Upsilon(x), \varphi(x) > = \int_0^\infty \varphi(x)\, dx$$

Il est clair que la distribution $\{\Upsilon'(x)\}$ est la distribution identiquement nulle, c'est-à-dire que $<\{\Upsilon'(x)\}, \varphi(x)> = 0$ quel que soit $\varphi \in \mathcal{D}$. Par contre, on peut montrer que :

Théorème VIII.- La dérivée de la distribution $\Upsilon(x)$ de Heaviside est la distribution de Dirac δ (définie par (6)) :

$$\frac{d\,\Upsilon(x)}{dx} = \delta \qquad (16)$$

Définition.- On appelle translatée de L de la distribution T_x la distribution définie par :

$$< T_{x-L}, \varphi(x) > = < T_x, \varphi(x+L) >$$

On en tire immédiatement le théorème suivant :

Théorème IX.- Les opérations translation et dérivation sont permutables.

On en déduit évidemment que la dérivée $\Upsilon_{x^i}(x)$ fonction de répartition d'une variable aléatoire, presque certainement égale à x^i n'est autre que δ_{x^i} translatée de la distribution de Dirac définie par (7).

On conçoit donc que la fonction de Heaviside et la distribution de Dirac sont destinées à jouer un rôle capital en calcul des probabilités. Cette importance est encore accrue par le théorème suivant.

Théorème X.- Si la fonction $F(x)$ a une dérivée $F'(x)$ pour $x \neq x^1, x^2, \ldots x^n$ et si on pose $p_i = F(x^i + 0) - F(x^i - 0)$, alors la dérivée de la distribution F est la distribution

$$F'(x) = \{F'(x)\} + \sum p_i\, \delta_{x^i} \qquad (17)$$

2) Cas de l'espace à n dimensions

Définition.- On appelle dérivée partielle de la distribution notée T_x ou $T_{x_1 \ldots x_n}$ par rapport à x_i la fonctionnelle $\frac{\partial T}{\partial x_i}$ définie par :

$$< \frac{\partial T}{\partial x_i}, \varphi > = - < T, \frac{\partial \varphi}{\partial x_i} > \qquad (18)$$

Evidemment, on peut répéter sur cette dérivée partielle tous les théorèmes de la section précédente.

Naturellement, on peut considérer ensuite les dérivées successives si par exemple nous considérons l'opérateur de dérivation $\frac{\partial^{r_1}}{(\partial x_1)^{r_1}} \cdots \frac{\partial^{r_n}}{(\partial x_n)^{r_n}}$ alors on a

$$< \frac{\partial^{r_1}}{(\partial x_1)^{r_1}} \cdots \frac{\partial^{r_n}}{(\partial x_n)^{r_n}} T, \varphi > = (-1)^{r_1+\cdots+r_n} < T, \frac{\partial^{r_1}}{(\partial x_1)^{r_1}} \cdots \frac{\partial^{r_n}}{(\partial x_n)^{r_n}} \varphi > \quad (19)$$

Parmi tous les opérateurs de dérivation, il en est un destiné à jouer un rôle primordial en calcul des probabilités, c'est l'opérateur $\frac{\partial}{\partial x_1} \cdots \frac{\partial}{\partial x_\ell} \cdots \frac{\partial}{\partial x_n}$ que nous désignerons dans la suite par $\frac{\partial}{\partial x}$ ou par $\mathcal{D}$. On a évidemment :

$$< \frac{\partial}{\partial x} T, \varphi > = (-1)^n < T, \frac{\partial \varphi}{\partial x} > \quad (20)$$

formule qui généralise (12).

Il est clair que :

<u>Théorème XI.- Toute distribution est indéfiniment dérivable et que l'on peut intervertir l'ordre des dérivations</u> (1).

On généralise aisément à l'espace à n dimensions la notion de fonction de Heaviside : c'est la fonction notée $\Upsilon(x)$ ou $\Upsilon(x_1 \ldots x_n)$ qui est égale à la fonction de répartition d'une variable aléatoire presque certainement nulle.

On généralise aussi la distribution de Dirac et on appelle $\delta_{(x)}$ ou $\delta_{(x^1 \ldots x^n)}$ la distribution telle que $< \delta_{x_1 \ldots x_n}, \varphi > = \varphi(0, \ldots, 0$ et l'on aura :

$$\frac{\partial^n}{\partial x_1 \ldots \partial x_n} \Upsilon(x_1, \ldots x_n) = \delta_{x_1 \ldots x_n}. \quad (21)$$

ou sous forme condensée :

$$\frac{\partial \Upsilon(x)}{\partial x} = \delta_{(x)} \quad (21)$$

La translatée sera toujours définie par

$$< T_{x-L}, \varphi(x) > = < T_x, \varphi(x+L) >$$

$$= < T_{x_1 \ldots x_n}, \varphi(x_1 + x_1^1, x_2 + x_2^1 \ldots, x_n + x_n^1) >$$

et nous noterons la translatée de δ par $\delta_{x_1^1 \ldots x_n^1}$ ou plus brièvement par δ_{x^1}, on aura donc

$$< \delta_{x^1}, \varphi > = \varphi(x_1^1, \ldots x_n^1)$$

---

(1) Une fonction f (x) localement sommable apparait ainsi comme indéfiniment dérivable au sens des distributions. Mais sa dérivée sera une distribution qui ne pourra être identifiée à une fonction que si f est suffisamment régulière; c'est ce qui fait que les fonctions dérivées de f au sens usuel peuvent être telles que $\frac{\partial^2 f}{\partial x_1 \partial x_2} \neq \frac{\partial^2 f}{\partial x_2 \partial x_1}$ alors que pour les distributions dérivées on a toujours $\frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial^2 f}{\partial x_2 \partial x_1}$

3) Support d'une distribution

Définition : On appelle support de la distribution T_x le complémentaire du plus grand ouvert Ω tel que pour tout $\varphi \in \mathcal{D}$ ayant son support dans Ω, on ait $< T_x, \varphi > = 0$

En particulier on peut montrer que le support d'une distribution de probabilité est le complémentaire du plus grand ouvert tel que :

$$\Pr (X \in E) = 0 \quad \text{pour tout} \quad E \subset \Omega$$

Exemples :

α) Si f(x) est une fonction continue, la distribution f(x) aura pour support un ensemble fermé F qui est l'adhérence de l'ensemble des points x tels que $f(x) \neq 0$ ($x \in R^1$ ou de R^n), c'est-à-dire qu'il coïncidera avec le support de f considérée comme fonction.

β) le support de $\delta_{x_1, \ldots x_n}$ est l'origine

γ) dans R^1 le support de la distribution $\sum p_i \, \delta_{x^i}$ est constitué par les seuls points $x = x^i$.

Théorème XII. - Le support S de la distribution dérivée $\frac{dT}{dx}$ (ou $\frac{\partial^{n_1}}{\partial x_1^{p_1}} \cdots \frac{\partial^{p_n}}{\partial x_n^{p_n}} T$) est contenu dans le support de la distribution T.

Ces deux supports ne sont d'ailleurs pas nécessairement identiques par exemple le support de la distribution de Heaviside est l'intervalle $0, +\infty$ alors que le support de la distribution de Dirac δ se réduit à l'origine.

b) Thécrèmes relatifs aux distributions de probabilité

1) Cas de l'espace à une dimension

Nous allons prouver le théorème suivant :

Théorème XIII. - F(x) étant une certaine fonction de répartition, la dérivée de la distribution F(x) est une distribution de probabilité. Réciproquement toute distribution de probabilité est la dérivée d'une certaine fonction de répartition.

α) En effet, si F(x) est une fonction de répartition, $\frac{d\,F(x)}{dx}$ est une distribution de probabilité.

1°- la dérivée $\frac{dF}{dx}$ de la distribution F est une distribution (théorème VI)

2°- cette distribution est telle que :

$$< \frac{dF}{dx}, \varphi > = - < F, \frac{d\varphi}{dx} >$$

$$= - \int_{-\infty}^{+\infty} F(x) \frac{d\varphi}{dx} dx \qquad (22)$$

Mais comme par hypothèse le support de φ est un intervalle fermé ab (22) peut s'écrire :

$$< \frac{dF}{dx}, \varphi > = - \int_a^b F(x) \frac{d\varphi}{dx} dx \qquad (23)$$

ab étant un intervalle borné et φ étant une fonction continue, les conditions pour que s'applique le théorème d'intégration par parties de Saks (1) sont remplies et on a :

$$< \frac{dF}{dx}, \varphi > = - \left\{ F(b+)\varphi(b+) - F(a-)\varphi(a-) - \int_a^b \varphi(x)\, d\, F(x) \right\}$$

$$= \int_a^b \varphi(x)\, d\, F(x) = \int_{-\infty}^{+\infty} \varphi(x)\, d\, F(x) \qquad (24)$$

et sous la forme (24) il est évident que si $\varphi \in \mathcal{D}$ est positive $< \frac{dF}{dx}, \varphi >$ le sera aussi.

3°- Si la suite des φ_i tend vers la fonction continue $\chi(x)$ à support borné comme il est dit dans 1°, il est clair que $\int_{+\infty}^{+\infty} \varphi_i(x)\, d\, F(x)$ tendra vers $\int_{-\infty}^{+\infty} \chi(x)\, d\, F(x)$.

Si ensuite la suite des fonctions continues $\chi_i(x)$ converge presque partout vers la fonction $1_{a,b}$, il est clair que $\int_{-\infty}^{+\infty} \chi_i(x)\, d\, F(x)$ convergera vers $\int_a^b d\, F(x) = F(b) - F(a)$.

Enfin quand $b \to -\infty$ et $a \to -\infty$, $F(b) - F(a) \to 1$

donc : $$< \frac{dF}{dx}, 1_{-\infty,+\infty} > = 1$$

ce qui achève de démontrer la première partie du théorème.

β) Réciproquement, montrons que toute distribution de probabilité est la dérivée d'une fonction de répartition.

Soit en effet T une distribution de probabilité. D'après le théorème IV, il existe une mesure positive $\mu(x)$ telle que :

$$< T, \varphi > = \int_a^b \varphi(x)\, d\mu(x) \qquad (25)$$

$$= - \int_a^b \mu \frac{d\varphi}{dx} dx \quad \text{d'après le théorème de Saks}$$

$$= - < \mu, \frac{d\varphi}{dx} >$$

$$= < \frac{d\mu}{dx}, \varphi >$$

donc : $$T = \frac{d\mu}{dx} \qquad (27)$$

---

(1) Saks. Théorie de l'intégrale (p. 102 de l'édition anglaise).

En outre de (25) et du fait que $\langle T, 1_{-\infty,+\infty}\rangle = 1$ on déduit que $\int_{-\infty}^{+\infty} d\mu(x) = 1$, ce qui achève de démontrer le théorème puisque μ est déterminé à une constante additive près.

Cas de l'espace à n dimensions

Théorème XIV.- $F(x) = F(x_1, x_2, \ldots, x_n)$ étant une certaine fonction de répartition, la "dérivée" $\frac{\partial}{\partial x} = \frac{\partial}{\partial x_1}\frac{\partial}{\partial x_2}\cdots\frac{\partial}{\partial x_n}$ de la distribution $F(x)$ est une distribution de probabilité dans R^n.

Réciproquement, toute distribution de probabilité T dans R^n est la "dérivée" d'une certaine fonction de répartition.

Ce théorème se démontre exactement de la même manière que le précédent.

Ceci posé, nous avons maintenant les éléments de base nécessaires pour traiter les applications que nous avons en vue. A cet effet, nous modifierons un peu nos notations et appellerons désormais $F(x,y)$ la fonction de répartition du couple aléatoire XY.

5. - ÉTUDE DES COUPLES ALÉATOIRES DONT LES DISTRIBUTIONS MARGINALES SONT CONNUES

Considérons donc un couple aléatoire XY dont la distribution T_{xy} est inconnue, mais dont on connait les distributions marginales R_x et S_y.

Divers problèmes se posent à propos de ce couple.

a) Quelle relation y a-t-il entre R_x, S_y et R_{xy} ?

On a évidemment

Théorème XV.- $R_x = \langle T_{xy}, 1_{-\infty,+\infty}\rangle$ (28)

En effet, la fonction de répartition $F(x,y)$ du couple XY est donnée par :

$$F(x,y) = \langle T_{\xi\eta}, 1_{(-\infty,+\infty),(x,y)}\rangle \qquad (29)$$

la fonction de répartition marginale de la v.a. X vaut donc :

$$F(x,\infty) = \langle T_{\xi\eta}, 1_{(-\infty,+\infty),(x,+\infty)}\rangle \qquad (30)$$

$$= \langle\langle T_{\xi\eta}, 1_{-\infty,+\infty}\rangle 1_{-\infty,x}\rangle \qquad (31)$$

Mais on a aussi :

$$F(x,\infty) = \langle R_\xi, 1_{-\infty,x}\rangle$$

ce qui démontre (28).

b) Distributions T_{xy} remarquables

Ceci posé Maurice Fréchet [1] remarque que d'une part on a toujours :

$$F(x,y) \leqslant \text{Min} \{ F(x,\infty), F(\infty,y) \} \tag{32}$$

et d'autre part :

$$\begin{aligned} F(x,y) &\geqslant \text{Max} \{ 0, 1 - [1 - F(x,\infty)] - [1 - F(\infty,y)] \} \\ &\geqslant \text{Max} \{ 0, F(x,\infty) + F(\infty,y) - 1 \} \end{aligned} \tag{33}$$

et il a montré que les fonctions :

$$F^{(1)}(x,y) = \text{Min} \{ F(x,\infty), F(\infty,y) \} \tag{34}$$

$$F^{(2)}(x,y) = \text{Max} \{ 0, F(x,\infty) + F(\infty,y) - 1 \} \tag{35}$$

étaient effectivement des fonctions de répartition et définissaient dans le cas où X est une v.a. discrète et dans le cas où les fonctions de répartition marginales sont continues, ce qu'on peut appeler des liaisons statistiquement non décroissantes et statistiquement non croissantes. Nous nous proposons ici de donner une autre démonstration des théorèmes de Fréchet, démonstration qui a le mérite de préciser ce qu'il faut entendre par liaison statistiquement croissante (décroissante) dans le cas général. Considérons donc la fonction $F^{(1)}(x,y)$ définie par (3) et cherchons s'il existe une distribution de probabilité telle que :

$$< T^{(1)}_{\xi\eta}, 1_{(-\infty,-\infty),xy} > = F^{(1)}(x,y) \tag{36}$$

la démonstration comprend 3 points.

1°) Il existe une distribution vérifiant (36) :

c'est $$T^{(1)}_{xy} = \frac{\partial F^{(1)}(x,y)}{\partial x\, \partial y}$$

2°) Cette distribution est positive, en effet :

$$< \frac{\partial^2 F^{(1)}}{\partial x\, \partial y}, \varphi > = < F^{(1)}, \frac{\partial^2 \varphi}{\partial x\, \partial y} >$$

$$= \iint F^{(1)}(x,y) \frac{\partial^2 \varphi}{\partial x\, \partial y}\, dx\, dy$$

Or $F^{(1)}$ étant à variation bornée, on peut appliquer le théorème de Saks, il vient :

$$< \frac{\partial^2 F^{(1)}}{\partial x\, \partial y}, \varphi > = \iint \varphi\, d\, F^{(1)}(x,y)$$

et sous cette forme, on voit que si $\varphi \in \mathcal{D}$ est positive $\frac{\partial^2 F^{(1)}}{\partial x\, \partial y}$ est positive puisque l'élément d'intégration $d\, F^{(1)}$ l'est (comme on le vérifie sans peine)

(1) Maurice Fréchet. Sur les tableaux de corrélation dont les marges sont données. Ann. Univ. Lyon, 9, 1951, p. 53-7.

3°) On a : $< T^{(1)}_{\xi\eta}, 1_{(-\infty,-\infty),(+\infty+\infty)} > = F^{(1)}(+\infty,+\infty)$

$$= \mathrm{Min}\left\{ F(\infty,\infty)\ ,\ F(\infty,\infty)\right\}$$

$$= 1$$

Nous avons donc démontré que $T^{(1)}_{xy}$ est bien une distribution de probabilité. Nous nous proposons maintenant de chercher le support de cette distribution.

A cet effet, remarquons que le plan peut être séparé en 2 régions E_1 et E_2 telles que l'on ait :

pour E_1 $\qquad F(x,\infty) \geqslant F(\infty,y)$

pour E_2 $\qquad F(x,\infty) < F(\infty,y)$

Soient E'_1 et E'_2 les plus grands ouverts contenus dans E_1 et E_2, $\mathscr{C}_1$ leur frontière. Nous nous proposons de montrer que le support de $T^{(1)}_{xy}$ est tout entier contenu dans $\mathscr{C}_1$. Pour cela, il suffit de montrer que quelle que soit $\varphi \in \mathscr{D}$ et ayant son support dans E'_1 ou dans E'_2 on a $< T^{(1)}_{xy}, \varphi > = 0$. Supposons d'abord que φ a son support dans E'_1, alors on a :

$$< T^{(1)}_{xy}, \varphi > = < \frac{\partial^2 F^{(1)}}{\partial x\, \partial y}, \varphi > = - < \frac{\partial F^{(1)}}{\partial x}, \frac{\partial \varphi}{\partial y} >$$

mais sur E'_1, $F^{(1)}(x,y) = F(\infty,y)$ la dérivée $\frac{\partial F^{(1)}}{\partial x}$ au sens des fonctions existe donc et est nulle, la distribution pourra donc être identifiée à la fonction nulle, donc $< T_{xy}, \varphi > = 0$. On verrait de même que si φ a son support sur E'_1

$$F^{(1)}(x,y) = F(x,\infty)$$

et

$$< T^{(1)}_{xy}, \varphi > = - < \frac{\partial F^{(1)}}{\partial y}, \frac{\partial \varphi}{\partial x} > = 0$$

ce qui démontre le théorème.

Il est aisé par ailleurs de voir que la courbe $\mathscr{C}_1$ peut être représentée par une fonction non décroissante de x.

Pour le prouver, il suffit de montrer que si $x_0 y_0$ est un point de E_1, $x\, y_0$ est aussi un point de E_1 pour tout $x > x_0$ autrement dit que si on a :

$$F(x_0,\infty) \geqslant F(\infty,y_0)$$

on a aussi $\qquad F(x,\infty) \geqslant F(\infty,y_0) \qquad$ pour tout $\quad x > x_0$

ce qui est évident.

Généralement d'ailleurs le support de $T^{(1)}_{xy}$ ne coincide pas avec $\mathscr{C}_1$. Si par exemple les v.a. marginales sont discontinues et susceptibles de prendre seulement les valeurs $x_i\ y_j$ on aura :

$$R_x = \sum_i \delta_{x_i}\ p_i$$

$$S_y = \sum_j \delta_{y_j}\ p_j$$

$$T_{xy} = \sum_i \sum_j \delta_{x_i y_j}\ p_{ij}$$

où les p_{ij} représentent les sauts aux points $x^i\, y^j$ de la fonction : $F^{(1)}(x,y) = \text{Min}\left\{F(x,\infty),\ F(\infty,y)\right\}$

Dans ces conditions, le support de T_{xy} est évidemment constitué par les seuls points $x^i\, y^j$ de $\mathscr{C}_1$ pour lesquels $p_{ij} \neq 0$.

Cette dernière remarque nous conduit à essayer de déterminer exactement quel est le support de $T_{xy}^{(1)}$ connaissant R_x et S_y. Soient A le support de R_x et B le support de S_y, on peut démontrer le lemme suivant :

Lemme. - Le support de $T_{xy}^{(1)}$ est constitué par l'ensemble des points de C_1 qui appartiennent à l'ensemble $A \times B$.

Par ailleurs, il est aisé de voir qu'il existe une distribution de probabilité $T_{xy}^{(2)}$ telle que (35) soit vérifiée et que cette distribution a son support tout entier contenu sur la courbe $\mathscr{C}_2$ définie par une certaine fonction non croissante (Pour le voir on peut, soit reprendre pas à pas la démonstration précédente, soit plus simplement se ramener au cas précédent en remarquant que (35) peut s'écrire :

$$F(\infty,y) - F(x,y) = \text{Min}\left\{F(\infty,y),\ 1 - F(x,\infty)\right\}$$

ou

$$\Pr\left\{X > x,\ Y \leq y\right\} = \text{Min}\left\{\Pr(y \leq Y),\ \Pr(X > x)\right\}$$

évidemment on ne pourra avoir $T_{xy}^{(1)} = T_{xy}^{(2)}$ que si ou bien il existe une valeur x_0 de x telle que $R_x = \delta_{x_0}$ ou bien il existe une valeur y_0 de y telle que $S_y = \delta_{y_0}$, c'est-à-dire si l'une des variables aléatoires XY est presque sûrement constante.

Les résultats de ce paragraphe sont résumés dans le théorème suivant :

Théorème XVI.-

1°- Il existe en général une infinité de couples aléatoires XY de distribution de probabilité T_{xy}, admettant des distributions marginales R_x, S_y connues (Il n'en serait autrement que si on peut trouver x_0 ou y_0 telles que l'on ait ou bien $R_x = \delta_{x_0}$ ou bien $S_y = \delta_{y_0}$).

2°- Si nous ordonnons l'ensemble E des T_{xy} en convenant de dire que $T'_{xy} < T''_{xy}$ si $F'(x,y) \leq F''(x,y)$ quels que soient x et y l'ensemble E a un élément maximum $T_{xy}^{(1)}$ et un élément minimum $T_{xy}^{(2)}$.

3°- Le support de $T_{xy}^{(1)}$ est tout entier contenu sur une courbe $\mathscr{C}_1$ qui peut être représentée par une fonction non décroissante et de même le support de $T_{xy}^{(2)}$ est tout entier contenu sur une courbe $\mathscr{C}_2$ qui peut être représentée par une fonction non croissante.

6. - APPLICATION AU PROBLÈME DE BERTAUT

Dans une note aux Comptes Rendus (1) J. Bertaut dit ceci (textuellement aux notations près) :

"Définitions : Soient n variables stochastiques $X_1, \ldots, X_n$ et soit $dx = dx^1\, dx^2\, dx^n$ la probabilité élémentaire définie de telle sorte que :

$$\int_0^1 dx = \int_0^1 \prod dx = \int_0^1 \ldots \int_0^1 dx_1 \ldots dx_n = 1$$

(1) J. Bertaut. Sur la probabilité de valeurs de fonctions. Applications à la cristallographie. C.R. 10 janvier 1955, p. 152, t. 240.

Soit ∂ (z-a) la fonction de Dirac nulle pour $z \neq a$ infinie pour z = a et telle que

$$\int_{-\infty}^{+\infty} \partial(z-a)\, dz = 1 \qquad (37)$$

Soit enfin une fonction $\psi(x_1, \ldots, x_n)$ prenant toutes ses valeurs dans l'intervalle $0 < x_j < 1$ $(j = 1, \ldots n)$, elle peut être zéro à l'extérieur de l'intervalle ou encore être périodique.

Théorème.- La probabilité P(y) dy pour que la fonction $\psi(X_1,\ldots,X_n)$ des n variables indépendantes X_j prenne des valeurs comprises entre y et y + dy est telle que :

$$P(y) = \int_0^1 \partial(\psi - y) \prod_{j=1}^{n} dx_j \qquad (38)$$

Outre que nous n'admettons pas l'existence de la fonction de Dirac l'intégrale (37) étant évidemment nulle. Nous voyons mal comment on peut en appliquant (38) trouver une évaluation correcte de la densité de probabilité P(y). Monsieur Bertaut ne dit pas comment il s'y est pris pour démontrer son soi-disant théorème. Il est probable qu'il a remarqué que : si le couple aléatoire XY a une densité de probabilité, si de plus X est uniformément réparti sur le pavé 0 1, et si enfin Y est en corrélation dure (1) par rapport à X, la densité de probabilité liée $f_x(y)$ est de la forme $\alpha\left[y - \psi(\vec{x})\right]$ et la densité de probabilité de y est :

$$f(y) = \int_0^1 \alpha\left[y - \psi(x)\right] \prod_{j=1}^{n} dx_j \qquad (39)$$

et il est probable que de (39) il a inféré (38).

Il est aisé de mettre (38) sous une forme rigoureuse. On a :

Théorème XVII.- La distribution T_y de la variable aléatoire : $Y = \psi(X_1 \ldots X_n)$ est telle que $\langle T_y, \varphi(y) \rangle = \langle\langle \delta_\psi, \varphi(y) \rangle 1_{0,1} \rangle$ où $1_{0,1}$ désigne la fonction identique à 1 sur le pavé $(0, \ldots, 0), (1, \ldots 1)$.

On peut démontrer ce théorème, soit en passant par les fonctions de répartition en utilisant la formule :

$$F(\infty, y) = \int \cdots \int F_x(y)\, dx$$

soit plus simplement en utilisant le théorème de Schwartz sur la convergence d'une suite de distributions $f_i(y)$ vers la distribution δ.

(1) Pour la définition de la notion de corrélation dure, voir S. Bernstein. Fondements géométriques de la théorie des corrélations. Metron 1928.

PRODUIT DE FONCTIONS CARACTÉRISTIQUES

ET INDÉPENDANCE DE VARIABLES ALÉATOIRES

par

M. GIRAULT

OBJET DE L'ÉTUDE

On considère les variables aléatoires X, Y et Z = X+Y et soient $\varphi_x(t)$, $\varphi_y(t)$, $\varphi_z(t)$ les fonctions caractéristiques associées. On sait que si X et Y sont indépendantes : $\varphi_z(t) = \varphi_x(t)\ \varphi_y(t)$. Nous voudrions attirer l'attention sur le fait que la réciproque n'est pas exacte ; et plus généralement sur le fait que les lois de répartition de X ; de Y (lois marginales) et de (X + Y) ne déterminent pas complètement la loi du couple (X, Y).

On notera en particulier que si X, Y et Z suivent des lois du χ^2 à n_1, n_2 et $(n_1 + n_2)$ degrés de liberté, cela n'entraine nullement l'indépendance de X et de Y.

DÉVELOPPEMENT

Soit donc le couple aléatoire (X, Y) et désignons par $F_{(x)}$ et $G_{(y)}$ les fonctions de répartition marginales de X et de Y.

Si les variables X et Y sont indépendantes, la fonction de répartition $H_{(z)}$ de Z = X + Y est obtenue par produit de composition de $F_{(x)}$ et de $G_{(y)}$.

(1) $$H(z) = F(x) * G(y) \quad \text{soit} \quad H(z) = \int F(z-y)\, d\, G(y)$$

D'autre part un résultat maintenant classique [1] nous apprend qu'il y a équivalence entre la relation (1) et la relation (2) suivante pour les fonctions caractéristiques :

(2) $$\varphi_z(t) = \varphi_x(t)\ \varphi_y(t)$$

Si $\Phi_{(r,s)} = E\left[e^{i(rx+sy)}\right]$ est la f.c. du couple (X, Y), la relation (1) entraine (3) :

(3) $$\Phi\,(t,t) = \Phi\,(t,o) .\ \Phi\,(o,t)$$

---

(1) Voir H. ROBBINS. Mixture of distributions. Annals of Math. Stat. Vol XIX, n°3 Sept. 1948, p. 365 ; et M. GIRAULT. Les fonctions caractéristiques et leurs transformations. Publications I.S.U.P. 1955.

puisque $\varphi_z(t) = \Phi(t,t)$ et que les f. c. des distributions marginales sont $\Phi(r,o)$ et $\Phi(o,s)$

Réciproquement, la relation (3) entraine (1) mais elle n'implique pas l'indépendance de X et de Y.

La condition nécessaire et suffisante d'indépendance de X et de Y relativement à $\Phi(r,s)$ est

$$\Phi(r,s) = \Phi(r,o).\ \Phi(o,s) \tag{4}$$

et cela pour toutes les valeurs de r et de s. Or la relation (3) indique seulement que (4) est satisfaite sur la droite r = s et non dans tout le plan. Il n'y a donc pas nécessairement indépendance stochastique.

En d'autres termes, on peut avoir la relation (1) pour certaines distributions de variables liées.

En effet, partons d'une répartition de v.a. indépendantes X et Y ; les fonctions de répartition de X de Y et de X + Y réalisent

$$H = F * G$$

Ajoutons une "pseudo répartition" de masses positives et négatives réalisant les conditions suivantes :

$$\text{I} \left\{ \begin{array}{ll} \text{masse nulle sur tout intervalle} & \Delta x \\ \text{"} \quad \text{"} \quad \text{"} \quad \text{"} \quad \text{"} & \Delta y \\ \text{"} \quad \text{"} \quad \text{"} \quad \text{"} \quad \text{"} & \Delta (x+y) \end{array} \right.$$

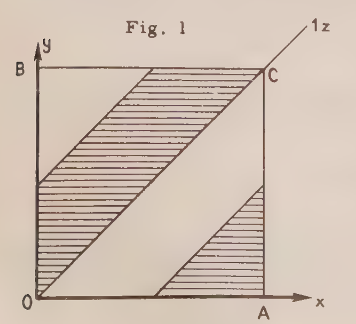
De plus, on imposera à la nouvelle répartition d'être partout à mesure positive.

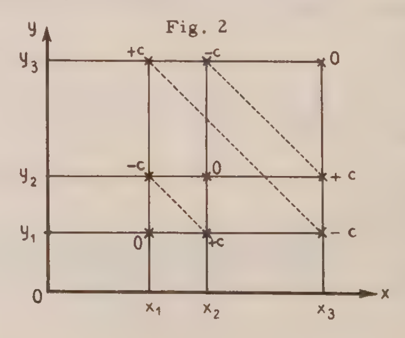
Il est bien clair que la nouvelle répartition admet encore F, G et H pour fonctions de répartition de X de Y et de (X + Y) et par suite on a encore : $H = F * G$ d'où $\varphi_z(t) = \varphi_x(t)\ \varphi_y(t)$

Il est facile de montrer que réciproquement, si une distribution réalise $H = F * G$, elle est la somme d'une distribution de v.a. indépendantes et d'une distribution de masses réelles réalisant les conditions (I).

Des exemples ont été donnés par H. Cramer (Mathematical Methods of Statistics 1946. Exercice p. 317, n° 2) et H. Robbins (Annals of Math. Stat. Sept. 1948). Dans ces exemples les conditions indiquées ici sont réalisées très simplement.

Voici à titre d'illustration l'exemple de H. Robbins :





Répartition uniforme de densité 2 dans les parties hachurées du carré (fig. I).

Les répartitions de X de Y et de Z sont les mêmes que celles données par une répartition uniforme dans tout le carré.

Ajoutons que la transformation indiquée par les conditions I est toujours possible dès que les répartitions de X et de Y ont au moins trois valeurs possibles : x , x_2, x_3, y , y_2, y_3, réalisant :

$$x_2 - x_1 = y_2 - y_1$$

$$x_3 - x_2 = y_3 - y_2$$

A partir de la répartition indépendante, on ajoutera aux points $(x_i y_j)$ les masses o, +c ou -c conformément à la figure II.

où $$c \leqslant \min. \text{ prob} \left\{ \begin{matrix} X = x_i \\ Y = y_j \end{matrix} \right\} \qquad i, j = 1, 2 \text{ ou } 3$$

Si la distribution est continue, les points $(x_i y_j)$ seront remplacés par des éléments dx dy et c par une densité.

En résumé les relations (1) ou (2) entrainent l'indépendance dans le seul cas où X et Y ne peuvent pas prendre l'un et l'autre au moins 3 valeurs ayant deux à deux des différences égales.

DES POSITIONS TYPIQUES D'UNE VARIABLE ALÉATOIRE

par

Saad K. NASR

1. - INTRODUCTION

Une définition de la moyenne d'une variable aléatoire située dans un espace de Banach a été donnée par Mourier (9). D'autre part, une seconde définition de cette moyenne a été exprimée dans Espace métrique par Doss (2). On sait (10) qu'une moyenne dans le sens de Mourier, si elle existe, est aussi une moyenne dans le sens de Doss. Fréchet (3), généralisa la définition de la moyenne de Doss en vue de définir les positions typiques d'une variable aléatoire située dans un espace métrique (ses positions : moyenne, centrale, équiprobable, dominante, etc...).

Nous définirons sous la définition -2-, dans les conditions les plus générales, les positions typiques d'une variable aléatoire quelconque de telle manière qu'une position typique existant dans le sens de Fréchet, sera aussi dans notre sens une position typique (proposition 2).

2. -

Soit $(f_i)_{i \in I}$ une famille de fonctions numériques (finies) définie sur un ensemble $\mathfrak{X}$, telle que pour deux éléments différents x et y appartenant à l'ensemble, il existe un $i \in I$ pour lequel $f_i(x) \neq f_i(y)$. On définit alors une structure uniforme séparée ((1), Ch. IX, parag.§1) sur $\mathfrak{X}$ par la famille (g_i) des écarts $g_i(x,y) = | f_i(x) - f_i(y) |$. Inversement, étant donnée une structure uniforme séparée sur $\mathfrak{X}$, une famille d'écarts est alors définie sur $\mathfrak{X}$ ((1), Ch. IX, parag.§1).

2.1. Plus généralement étant donnée une variable aléatoire X située dans l'ensemble $\mathfrak{X}$, on supposera qu'une structure uniforme séparée est définie sur $\mathfrak{X}$ par la famille d'écarts finis $d_i(x,y)$ avec $i \in I$. La classe K des transformations

$$d_\alpha(x) = d_{(i,a)}(x) = d_i(x,a) \qquad (1)$$

définie dans $\mathfrak{X}$ pour toute paire d'éléments $\alpha \equiv (i,a)$ avec $i \in I$ et a élément fixe de $\mathfrak{X}$, est alors une classe séparative ((8), p. 29).

Définition (1)

LA CLASSE K DE TRANSFORMATIONS $d_\alpha(x)$, définie sous le paragraphe 2.1. est appelée classe séparative de transformations métriques.

On écrira par la suite T(N) pour la position typique d'un nombre aléatoire N en supposant que pour tout α, $d_\alpha(X)$ est un nombre aléatoire pour lequel une position typique $T(d_\alpha(X))$ existe.

Définition (2)

Un élément $\tau(x)$ est dit position typique (correspondant au symbole T) d'une variable aléatoire X située dans un ensemble $\mathfrak{X}$, s'il existe une classe séparative $K = \{ d_\alpha(x) \}$ avec $x \in \mathfrak{X}$ de transformations métriques, de telle manière que l'on ait

$$d_\alpha(\tau(X)) \leqq T(d_\alpha(X)) \quad ; \quad \tau(X) \in \mathfrak{X}$$

indépendamment de α.

2.2. Considérons la variable aléatoire X située dans un espace métrique $(\mathcal{D})$ défini par la distance d. Un élément $\bar{x} \in (\mathcal{D})$ est une moyenne dans le sens de Doss (2), si pour tout élément fixe $a \in (\mathcal{D})$ on a $d(\bar{x},a) \leqq \mathfrak{M}\, d(X,a)$ où $\mathfrak{M}$ indique la moyenne classique d'un nombre aléatoire. Or $K = (d_\alpha(x)) \equiv (d(x,a))$ étant une classe séparative de transformations métriques, on voit que toute moyenne dans le sens de Doss, est moyenne dans le sens de notre définition (2). D'où, plus généralement :

PROPOSITION I

TOUTE POSITION TYPIQUE D'UNE VARIABLE ALEATOIRE X DANS LE SENS DE FRECHET (3) EST UNE POSITION TYPIQUE DE X DANS LE SENS DE LA DEFINITION (2).

2.3. Afin de justifier la condition que la classe K soit nécessairement une classe séparative, nous donnerons l'exemple suivant :

Soit une classe K définie par le seul écart $g(x,y) = |f(x) - f(y)|$ où

$$f(x) = \begin{cases} -x & \text{pour} \quad x < 0 \\ 0 & \text{pour} \quad 0 \leqq x \leqq 2 \\ x-2 & \text{pour} \quad x > 2 \end{cases}$$

Pour une variable aléatoire X qui prend les valeurs 0, 1 et 2, avec une probabilité égale à 1/3 dans chaque cas, on a

$$\mathfrak{M}\, g(X,a) = f(a) \qquad \text{pour tout a réel.}$$

Autrement dit, la moyenne classique $\bar{x} = 1$ de la variable aléatoire X n'est pas l'unique solution de l'inégalité

$$g(\tau,a) \leqq \mathfrak{M}\, g(X,a)$$

puisque celle-ci est satisfaite par tout élément τ dans l'intervalle (0,2).

3. -

Considérons à présent la variable aléatoire X située dans l'espace de Banach $(\mathfrak{X})$. Un élément $\bar{x} \in (\mathfrak{X})$ est une moyenne dans le sens de Mourier (9) si l'égalité

$$x^*(\bar{x}) = \mathfrak{M}\,(x^*(X)) \qquad (1)$$

est satisfaite pour toute fonctionnelle linéaire continue x^* définie dans $(\mathfrak{X})$.

3.1. On sait que (10) toute moyenne dans le sens de Mourier est aussi moyenne dans le sens de Doss, d'où, d'après la Proposition I, elle l'est aussi dans le sens de la Définition (2). D'ailleurs ceci peut être prouvé directement comme suit :

PROPOSITION II

LA MOYENNE DE MOURIER D'UNE VARIABLE ALEATOIRE $X \in (\mathfrak{X})$, si elle existe, EST AUSSI UNE MOYENNE DANS LE SENS DE LA DEFINITION (2).

En effet, pour n'importe quelle paire d'éléments distincts x et y de $(\mathfrak{X})$, il existe ((6), p. 19) une fonctionnelle linéaire x^* telle que

$$|x^*(x-y)| = \| x-y \| > 0 \qquad (2)$$

On peut donc dire qu'il existe un écart $d^*(x,y)$ défini par

$$d^*(x,y) = |x^*(x) - x^*(y)| \qquad (3)$$

tel que l'on ait

$$d^*(x,y) = |x^*(x) - x^*(y)| = |x^*(x-y)| > 0$$

D'où la famille des écarts $\{d^*(x,y)\} = \{|x^*(x) - x^*(y)|\}$ où x^* appartient à l'espace $(\mathfrak{X}^*)$ conjugué ((6), p. 21) de $(\mathfrak{X})$, définit une classe de transformations métriques.

Si $\bar{x} \in (\mathfrak{X})$ est une moyenne de Mourier d'une variable aléatoire X, on aura de l'égalité (1)

$$x^*(\bar{x}) - x^*(a) = \mathfrak{M}\,(x^*(X) - x^*(a)) \qquad (4)$$

indépendamment de $a \in (\mathfrak{X})$.

De (3) et (4) on a

$$d^*(\bar{x},a) = |\,\mathfrak{M}\,(x^*(X) - x^*(a)\,| \leqslant \mathfrak{M}\, d^*(X,a) \qquad (5)$$

ce qui démontre la proposition

3.2. LA DEFINITION DE FRECHET DE LA MOYENNE D'UNE VARIABLE ALEATOIRE X SITUEE DANS L'ESPACE DE BANACH ((2), parag. 5)

Dans ce qui suit, E dénote l'ensemble d'évènements élémentaires ξ, d'après la terminologie de Kolmogoroff (7). Une variable aléatoire X est dite appartenir à la classe (I F), si quelque soit $\varepsilon > 0$, on peut trouver une décomposition $\Delta (= \Delta_{\varepsilon,X})$ de l'ensemble E en une suite dénombrable d'en-

sembles disjoints probabilisables $e_1, e_2, \ldots$ ($E = \bigcup_{r=1}^{\infty} (e_r)$) pour laquelle la condition suivante est satisfaite :

$$\left\{ p(e_r) = \text{Prob.}(e_r) = 0 \qquad \text{ou} \qquad \max_{\xi,\xi' \in e_r} \| {}_{\xi}x - {}_{\xi'}x \| < \varepsilon \right\}, r=1,2,\ldots \quad \text{(i)}$$

de telle façon que la série

$$\sigma_\varepsilon = \sum_{r=1}^{\infty} p(e_r) \| {}_{\xi_r}x \| \qquad , \quad \xi_r \in e_r, \qquad \text{(ii)}$$

soit convergente.

Si la somme σ_ε existe pour au moins une décomposition $\Delta_{\varepsilon,X}$, elle existe pour toute autre décomposition $\Delta_{\varepsilon',X}$, (2).

L'espace de Banach étant complet, quand $\varepsilon \longrightarrow 0$, la limite de

$$S_\varepsilon = \sum_{r=1}^{\infty} p(e_r)_{\xi_r} x \qquad , \quad \xi_r \in e_r, \qquad \text{(iii)}$$

existe indépendamment de $\Delta_{\varepsilon,X}$. Cette limite $I(X)$ est la moyenne (IF) de X.

Si l'espace ($\mathfrak{X}$) est un espace de Banach séparable, une telle décomposition Δ satisfaisant (i) est équivalente à l'affirmation que X est presque sûrement contenu dans un sous-ensemble séparable ($\mathfrak{X}_1$) d'un espace de Banach $\mathfrak{X}$.

On sait que si (10, Théorème 3) la variable aléatoire X de la classe (IF) est faiblement mesurable ((6), p. 36), la moyenne I(X) est aussi une moyenne dans le sens de Mourier. Il en découle de la proposition II la

PROPOSITION III

SI LA VARIABLE ALEATOIRE X DE LA CLASSE (I,F) EST FAIBLEMENT MESURABLE, LA MOYENNE I (X) EST AUSSI UNE MOYENNE DANS LE SENS DE LA DEFINITION (2).

4. -

Il est important de savoir si une position typique $\tau(X)$ d'une variable aléatoire X existant relativement à une classe K_1 dans le sens de la définition (2), existe de même relativement à une autre classe K_2, où K_1 et K_2 sont deux classes séparatives de transformations métriques, seules transformations prises en considération par nous.

4.1. L'exemple suivant montre que cela n'est pas toujours le cas.

Soit le nombre aléatoire X prenant les valeurs x_1 et x_2 ($x_1 \neq x_2$) avec les probabilités respectives p et q ($p+q = 1$). Considérons les deux classes K_1 et K_2 définies respectivement par les deux distances d_1 et d_2 avec

$$d_1(x,y) = |x-y|$$

$$\text{et} \quad d_2(x,y) = 0 \qquad \text{si } x = y$$

$$= 1 \qquad \text{si } x \neq y.$$

On a alors $d_2(\bar{x}, x_1) = 1$ où $\bar{x} = px_1 + qx_2$, la moyenne classique de X. Comme $\mathfrak{M}\, d_2(X, x_1) = q$, $d_2(\bar{x}, x_1) > \mathfrak{M}\, d_2(X, x_1)$. Ceci montre que la moyenne classique $\bar{x}$, qui est une moyenne dans le sens de la définition 2, relativement à la classe K_1 n'est pas une moyenne relative à la classe K_2.

En particulier, on remarque qu'une moyenne dans le sens de Doss relative à une certaine distance, n'est pas nécessairement une moyenne dans le même sens relative à n'importe quelle autre distance.

4.2. Etant donnée une famille $(f_i)_{i\in I}$ d'écarts sur $\mathfrak{X}$, soit g_H l'enveloppe supérieure de la famille $(f_i)_{i\in H}$, où H est une partie finie quelconque de I. La famille d'écarts (g_H) est équivalente ((1) Ch. IX, parag. 1) à la famille (f_i); on dit que c'est la famille d'écarts obtenue en saturant (f_i) et que (g_H) est une famille d'écarts saturée.

Définition (3)

Soient K_1 et K_2 deux classes différentes définies par les deux familles saturées $(f_i)_{i\in I}$ et $(g_x)_{x\in K}$ sur le même ensemble $\mathfrak{X}$. Si U_1 et U_2 sont les deux structures uniformes induites par (f_i) et (g_x) respectivement, on dit alors que K_1 est moins fine que K_2, si U_1 est moins fine que U_2. K_1 est dite strictement moins fine que K_2, si pour tout $i \in I$ il existe un $x \in \mathfrak{X}$ tel que $f_i \leqslant g_x$.

4.3. Soit $(\mathfrak{X})$ un système linéaire ((6) p. 9) sur lequel on définit les deux classes K_1 et K_2 comme suit :

$$K_1 = \{ \| x - a \| \},$$

$$K_2 = \{ d_i (x,a) \}_{i \in I} \qquad , a \in (\mathfrak{X}), \qquad (1)$$

où $\|x\|$ est la norme de x, (d_i) la famille d'écarts saturée telle que pour tout i, étant donnés des éléments x_1, x_2, y_1, y_2 de $(\mathfrak{X})$ on ait :

$$d_i(x_1+x_2,\ y_1+y_2) \leqslant d_i(x_1,y_1) + d_i(x_2,y_2)$$

$$d_i(kx,\ ky) \leqslant |k|\, d_i(x,y)$$

pour tout nombre réel k.

Remarque : Nous entendons par position typique d'une variable aléatoire celle admise dans la définition (2), sauf affirmation contraire.

PROPOSITION IV

SOIT X UNE VARIABLE ALEATOIRE SITUEE DANS UN SYSTEME LINEAIRE $(\mathfrak{X})$, SUR LEQUEL ON A DEFINI LES DEUX CLASSES K_1 ET K_2 INTRODUITES PAR LES RELATIONS (1) CI-HAUT. SI K_2 EST MOINS FINE QUE K_1 ET SI L'ON SUPPOSE QUE $\mathfrak{M}(d_\alpha(X))$ EXISTE INDEPENDAMMENT DE $\alpha \equiv (i,a)$, TOUTE MOYENNE (IF) DE X EXISTANT RELATIVEMENT A K_1 EST AUSSI UNE MOYENNE RELATIVE A K_2.

En effet la classe K_2 étant moins fine que K_1, pour tout écart d_i et tout $\varepsilon' > 0$ ((1) Ch. IX, parag. 1), il existe un $\varepsilon > 0$ tel que

$$\|x - y\| < \varepsilon \qquad \text{implique} \quad d_i(x,y) < \varepsilon' \qquad (2)$$

La variable aléatoire X étant de la classe (IF), on peut dire de (3.2.) que pour tout $\varepsilon > 0$, il existe une décomposition $\Delta_{\varepsilon,x}$ de E en une suite dénombrable d'ensembles disjoints probabilisables $e_1, e_2, \ldots$ $(E = \bigcup_{r=1}^{\infty} (e_r))$ pour laquelle la condition (i) de 3.2. est satisfaite.

d_i étant un écart, de (2) on obtient

$$\max_{\xi,\xi'\in e_r} |d_i({}_\xi x,a) - d_i({}_{\xi'} x,a)| < \max_{\xi,\xi'\in e_r} d_i({}_\xi x, {}_{\xi'} x) < \varepsilon' \quad , \qquad r = 1, 2, \ldots, \tag{3}$$

pour tout élément fixe a de $(\mathfrak{X})$.

De (3) et de (i) de 3.2. on voit que pour tout $\varepsilon' > 0$ il existe une décomposition $\Delta'_{\varepsilon' X} (= \Delta_{\varepsilon, X})$ de E en une suite dénombrable d'ensembles disjoints probabilisables $e_1, e_2, \ldots$ $(E = \bigcup_{r=1}^{\infty} (e_r))$; pour laquelle la condition

$$\{p(e_r) = 0 \quad \text{ou} \quad \max_{\xi,\xi'\in e_r} | d_i({}_\xi x,a) - d_i({}_{\xi'} x,a)| < \varepsilon' \}, r = 1, 2, \ldots \tag{4}$$

est satisfaite.

$$\text{Posons} \quad s_\varepsilon^{(n)} = \sum_{r=1}^{n} p(e_r)_{\xi_r} x \qquad \text{et} \qquad a_\varepsilon^{(n)} = a . \sum_{r=1}^{n} p(e_r) \tag{5}$$

Avec les conditions imposées à K_2 on obtient :

$$d_i(s_\varepsilon^{(n)}, a_\varepsilon^{(n)}) = d_i(\sum_{r=1}^{n} p(e_r)_{\xi_r} x, \sum_{r=1}^{n} p(e_r) . a) <$$
$$< \sum_{r=1}^{n} d_i(p(e_r)_{\xi_r} x, p(e_r).a) < \sum_{r=1}^{n} p(e_r) d_i({}_{\xi_r} x, a) \tag{6}$$

Comme la série (ii) de 3.2. est convergente, on a

$$\lim_{n\to\infty} s_\varepsilon^{(n)} = s_\varepsilon \qquad (\varepsilon \text{ fixe}) \tag{7}$$

On a de plus

$$\| a - a_\varepsilon^{(n)} \| = \| \sum_{r=n+1}^{\infty} p(e_r).a \| = \| a \| \sum_{r=n+1}^{\infty} p(e_r) \longrightarrow 0 \quad \text{as } n\to\infty \tag{8}$$

La topologie définie par K_2 étant moins fine que celle définie par K_1, on a des relations (7) et (8)

$$\lim_{n\to\infty} d_i(s_\varepsilon^{(n)}, s_\varepsilon) = \lim_{n\to\infty} d_i(a_\varepsilon^{(n)}, a) = 0 \tag{9}$$

Or d_i est un écart

$$| d_i(s_\varepsilon^{(n)}, a_\varepsilon^{(n)}) - d_i(s_\varepsilon, a) | \leq d_i(s_\varepsilon^{(n)}, s_\varepsilon) + d_i(a_\varepsilon^{(n)}, a) \tag{10}$$

Le passage à la limite $n \longrightarrow \infty$ dans l'inégalité (6), et de (9) et (10) il découle

$$d_i(s_\varepsilon, a) = \lim_{n\to\infty} d_i(s_\varepsilon^{(n)}, a_\varepsilon^{(n)}) < \lim_{n\to\infty} \sum_{r=1}^{n} p(e_r) d_i({}_{\xi_r} x, a)$$
$$= \sum_{r=1}^{\infty} p(e_r) d_i({}_{\xi_r} x, a) \tag{11}$$

Comme $\lim_{\varepsilon\to 0} s_\varepsilon = I(X)$ (de la relation (iii) de 3.2.), relative à la topologie définie par K_1, alors quand $\varepsilon' \longrightarrow 0$, (11) donne

$$d_i(I(X), a) \leq \lim_{\varepsilon'\to 0} \sum_{r=1}^{\infty} p(e_r) d_\alpha({}_{\xi_r} x) \qquad , \quad \xi_r \in e_r, \quad \alpha \equiv (i,a) \tag{12}$$

Or $\mathfrak{M}\, d_\alpha(X)$ existe pour tout $\alpha \equiv (i,a)$. $I(d_\alpha(X))$ existe ((2) parag. 5) et égale $\mathfrak{M}\, d_\alpha(X)$. Ceci entraine que le membre de droite de l'inégalité (11) existe et admette pour limite $\mathfrak{M}\, d_\alpha(X)$ quand $\varepsilon' \longrightarrow 0$.

c.q.f.d.

4.4. Nous donnons ici un exemple important de classes K_1 et K_2 satisfaisant les conditions (1) de 4.3.

PROPOSITION V

SOIT X UNE VARIABLE ALEATOIRE SITUEE DANS UN SYSTEME LINEAIRE $(\mathfrak{X})$ SUR LEQUEL DEUX CLASSES K_1 et K_2 SONT DEFINIES RESPECTIVEMENT PAR LES NORMES $\|x\|_1$, ET $\|x\|_2$. SI K_2 EST MOINS FINE QUE K_1, TOUTE MOYENNE (IF) DE X EXISTANT RELATIVEMENT A K_1 EST AUSSI MOYENNE RELATIVEMENT A K_2, AVEC LA SUPPOSITION QUE POUR TOUT $a \in (\mathfrak{X})$, $d_a(X) = \| X - a \|_2$ SOIT UN NOMBRE ALEATOIRE.

$\|x\|_2$ étant moins fine que $\|x\|_1$, il existe ((1), Ch. IX, parag. 3) un nombre réel $b > 0$ tel que

$$\|x\|_2 \leqslant b\, \|x\|_1 \qquad (1)$$

pour tout $x \in (\mathfrak{X})$. D'où pour tout $\varepsilon' > 0$, il existe $\varepsilon\ (=\frac{\varepsilon'}{b})$ tel que $\|x-y\|_1 < \varepsilon$ implique $\|x-y\|_2 < \varepsilon'$

En utilisant les mêmes notations que celles de la proposition IV, pour tout élément fixe a de $(\mathfrak{X})$, nous avons l'inégalité

$$\|s_\varepsilon - a\|_2 \leqslant \sum_{r=1}^{\infty} p(e_r) \|_{\xi_r} x - a\|_2 \quad , \qquad \xi_r \in e_r, \qquad (2)$$

qui est analogue à l'inégalité (11) de la proposition IV.

Comme $I(X)$ est une moyenne relative à K_1, alors $\mathfrak{M}\|X-a\|_1$ existe pour tout $a \in (\mathfrak{X})$. D'où la série $\sum_{r=1} p(e_r)\|_{\xi_r} x-a\|_1$, $\xi_r \in e_r$, existe et en conséquence l'existence du membre de droite de (2) découle de (1).

Comme $\lim_{\varepsilon\to 0}\|s_\varepsilon - I(X)\|_1 = 0$, $\lim_{\varepsilon\to 0}\| s_\varepsilon - I(X)\|_2 = 0$. En prenant les limites des deux membres de l'inégalité (2) quand $\varepsilon' \to 0$, on obtient

$$\|I(X) - a\|_2 \leqslant I(\|X - a\|_2) \qquad (3)$$

L'existence de la moyenne classique du nombre aléatoire $\|X-a\|_2$ résulte de l'existence de la moyenne classique du nombre aléatoire $\|X-a\|_1$ grâce à la relation (1) ; il en résulte de (3) la démonstration recherchée.

5. - LES POSITIONS TYPIQUES D'UNE VARIABLE ALÉATOIRE SITUÉE DANS L'ESPACE EUCLIDIEN R^n

Une position typique $\tau(X)$ d'une variable aléatoire X existant relativement à une classe K définie par une norme $\|x\|$ sera dite exister relativement à $\|x\|$.

Différentes normes sont définies dans l'espace euclidien ; nous choisirons la norme $\|x\|_1 = \sup_{1\leqslant i\leqslant n} |x_i|$, $x \equiv (x_1, \ldots, x_n) \in R^n$ pour déterminer relativement à lui quelques positions typiques d'une variable aléatoire X située dans R^n.

5.1. La moyenne : Soit $X \equiv (X_1, \ldots, X_n) \in R^n$ tel que $\overline{X_i} = \mathfrak{M} X_i$ existe pour tout $i = 1, 2, \ldots, n$, $\overline{X} \equiv (\overline{X}_1, \ldots \overline{X}_n)$ est une moyenne de X relativement à la norme $\|x\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2}$

On peut facilement prouver, soit directement, soit au moyen de la proposition V, que $\overline{X}$ est aussi une moyenne relativement à $\|x\|_1$.

5.2. Position centrale : Soit X une variable aléatoire stochastiquement bornée (3) située dans R^n, sur lequel une norme $\|x\|$, $(x \in R^n)$ est définie. Une position centrale $\gamma \in R^n$ de X, existe dans le sens de la définition (2) relativement à $\|x\|$, si $\|\gamma - a\| \leq \dot{N}$ avec $N = \|X - a\|$, indépendamment de $a \in R^n$, $\dot{N}$ étant la valeur centrale du nombre aléatoire N.

En général, relativement à $\|x\|_2$, une position centrale d'une variable aléatoire stochastiquement bornée $X \in R^n$, n'existe que ((3) p. 232) dans certains cas particuliers. Par contre, relativement à $\|x\|_1$, qui est <u>strictement moins fine</u> (définition 3) que $\|x\|_2$ si $X \equiv (X_1, \ldots X_n)$ est stochastiquement bornée, $\dot{X} \equiv (\dot{X}_1, \ldots \dot{X}_n)$ où $\dot{X}_i$ $(i = 1, \ldots n)$ est la valeur centrale du nombre aléatoire X_i, est aussi une position centrale de X. Car, $\dot{X}_i$ étant la position centrale de X_i, on a (5) :

$$|\dot{X}_i - a_i| \leq \dot{N}_i \qquad\qquad N_i = |X_i - a_i|$$

pour tout $i = 1, 2, \ldots, n$ avec a_i un nombre réel fixe quelconque.

Et si $a \equiv (a_1, \ldots, a_n)$ est un élément fixe quelconque de R^n, on a alors

$$0 \leq N_i \leq \|X - a\|_1 = N \qquad \text{d'où} \qquad \dot{N}_i \leq \dot{N}$$

Ceci entraine finalement que $|\dot{X}_i - a_i| \leq \dot{N}$ pour $i = 1, \ldots, n$, prouvant l'affirmation ci-haut.

On remarque qu'une position centrale de $X \in R^n$, si elle existe (3) relativement à $\|x\|_2$, existe aussi relativement à $\|x\|_1$, mais que l'inverse n'est pas nécessairement vrai, comme on peut le voir par les exemples particuliers donnés par Fréchet ((3) p. 232).

De plus, la position centrale $\dot{X} \equiv (\dot{X}_1, \ldots, \dot{X}_n)$ de $X \in R^n$, relativement à $\|x\|_1$ est unique. En effet, s'il existe une autre position centrale $\gamma \equiv (\gamma_1, \ldots, \gamma_n) \in R^n$ de X, $(\gamma \neq \dot{X})$, relativement à $\|x\|_1$, $\gamma_j \neq \dot{X}_j$ pour un certain entier j, $(1 \leq j \leq n)$. La variable aléatoire X étant stochastiquement bornée, on peut choisir un nombre réel c assez grand ($c > 2K$ si $\|x\|_1 < K$) de telle manière que

$$|{}_\xi x_i| \leq \|{}_\xi x\|_1 < K < |{}_\xi x_j - c|$$

pour tout $i \neq j$ et $\xi \in$ E (l'ensemble des évènements élémentaires).

Pour $a_0 \equiv (a_1, \ldots, a_n)$ avec $a_j = c$ et $a_i = 0$ $(i \neq j)$, on a

$$N_0 = \|X - a_0\|_1 = |X_j - c|$$

Comme γ est une position centrale de X, $|\gamma_j - c| \leq \|\gamma - a_0\|_1 \leq \dot{N}_0$, ou

$$|\gamma_j - c| \leq \dot{N}_j \qquad \text{où} \qquad N_j = |X_j - c| \qquad (1)$$

Ceci est contradictoire puisque $\dot{X}_j$ est la solution unique (5) de l'inégalité (1)

5.3. Position équiprobable : Rappelons d'abord, afin d'éviter toute confusion, qu'on appelle valeur équiprobable d'un nombre aléatoire N, tout nombre certain ν tel que l'on ait simultanément

$$\left\{\text{Prob}[N \leq \nu]\right\} \geq \frac{1}{2} \leq \text{Prob}[N \geq \nu]$$

On représente un tel point ν par la notation $\overline{\overline{N}}$.

Comme au paragraphe 5.2. on a vu que relativement à $\|x\|_2$, une position équiprobable ((3) p.235) d'une variable aléatoire $X \equiv (X_1, ..X_n) \in R^n$ n'existe pas en général, mais qu'elle n'existe que dans certains cas particuliers. Mais relativement à $\|x\|_1$, $\overline{\overline{X}} \equiv (\overline{\overline{X}}_1, ..., \overline{\overline{X}}_n)$ où $\overline{\overline{X}}_i$ $(i = 1, ..., n)$ est une valeur équiprobable du nombre aléatoire X_i, est une position équiprobable de X.

En effet, pour $N_i = |X_i - a_i|$ avec a_i = nombre fixe réel quelconque et $i = 1, ..., n$ on a (5)

$$|\overline{\overline{X}}_i - a_i| \leqslant \overline{\overline{N}}_i$$

et la validité de l'affirmation ci-haut se trouve établie à l'aide de la proposition suivante, puisque $N_i \leqslant N = \|x - a\|_1$.

PROPOSITION VI

SOIENT $\mu'(N)$ et $\mu''(N)$ LES VALEURS EQUIPROBABLES INFERIEURE ET SUPERIEURE D'UN NOMBRE ALEATOIRE N ((4), p. 45). SI $\{PROB[X \leqslant Y]\} = 1$, X ET Y ETANT DEUX NOMBRES ALEATOIRES, ON A

$$\mu'(X) \leqslant \mu'(Y) \quad \text{ET} \quad \mu''(X) \leqslant \mu''(Y)$$

En effet, si $\mu'(X) > \mu'(Y)$, comme $\{Prob[(X \leqslant Y)]\} = 1$, on a

$$\{Prob[Y > \mu'(Y)]\} \geqslant \{Prob[X > \mu'(Y)]\} \geqslant Prob[X > \mu'X) - \varepsilon]$$

où $2\varepsilon = \mu'(X) - \mu'(Y) > 0$.

D'où $\{Prob[Y > \mu'(Y)]\} > 1/2$

ce qui mène à une contradiction puisque $\mu'(Y)$ est valeur équiprobable inférieure de Y. De même, si $\mu''(X) > \mu''(Y)$, comme $\{Prob[(X \leqslant Y)]\} = 1$, on a

$$\{Prob[X \geqslant \mu''(X)]\} \leqslant \{Prob[Y \geqslant \mu''(X)]\} \leqslant Prob[Y > \mu''(Y) + \eta$$

où $2\eta = \mu''(X) - \mu''(Y) > 0$.

Donc $\{Prob[X \geqslant \mu''(X)]\} < 1/2$,

ce qui est aussi contradictoire, puisque $\mu''(X)$ est la valeur équiprobable supérieure de X.

BIBLIOGRAPHIE

(1) N. BOURBAKI : Les structures fondamentales de l'analyse. Tome 3. Topologie générale - Paris, Hermann, 1940.

(2) S. DOSS : Sur la moyenne d'un élément aléatoire dans un espace distancié (Bull. Sc. Math. 73, 1949, 48-72).

(3) M. FRECHET : Positions typiques d'un élément aléatoire de nature quelconque (Ann. Ec. Norm. Sup. 56-1948, 211-237)

(4) M. FRECHET : Recherches théoriques modernes sur le calcul des probabilités, Livre I, 2e édition (1950), Paris

(5) M. FRECHET : Une propriété des valeurs typiques d'un nombre aléatoire (Ann. Ec. Norm. (3) LXV - Fasc. 3)

(6) E. HILLE : Functional analysis and semi-groups (Amer. Math. Soc. Colloquium publications, vol. XXXI).

(7) A. KOLMOGOROFF : Grundbegriffe der Wahrscheinlichkeits-rechnung (Springer, Berlin, 1933)

(8) S. LEFSCHETZ : Algebraic topology (Amer. Math. Soc. Colloquim publications, vol. XXVII)

(9) E. MOURIER : Sur l'espérance mathématique d'un élément aléatoire dans un espace de Banach (C.R. Acad. Sc. Paris 229 -1949, 1300/1301).

(10) SAAD K. NASR : On random elements (Thesis) - Alexandrie, 1953.

DES DISTRIBUTIONS SYMÉTRIQUES A FONCTIONS CARACTÉRISTIQUES CONVEXES

par

Aurel WINTNER

Soit $\varphi(t)$ une fonction paire, définie sur la droite $-\infty < t < \infty$ et satisfaisant aux conditions suivantes :

(1) $$\varphi(t) \text{ est convexe (et continue) pour } 0 \leqq t < \infty$$

et $\varphi(0) = 1$, $\varphi(\infty) = 0$. La fonction $\varphi(t)$, évidemment, est monotone et non-décroissante pour $0 \leqq t < \infty$, et, suivant Jensen, elle a pour $0 < t < \infty$ des dérivées unilatérales finies, égales excepté sur un ensemble dénombrable de valeurs de t. De plus, $\varphi'(t)$ étant l'une de ces dérivées, $-\varphi'(t)$ est non-négative et non décroissante pour $0 < t < \infty$. Il est clair également que $\varphi(t)$ est absolument continue (de fait, elle satisfait à la condition de Lipschitz pour $\varepsilon < t < \infty$, si $\varepsilon > 0$). Puisque l'intégrale de $\varphi'(t)$, prise de 0 à ∞, est convergente, il s'ensuit d'après un lemme d'Abel que $t\varphi'(t) \to 0$ quand $t \to 0$ et quand $t \to \infty$.

Considérons la fonction

(2) $$\int_{+0}^{\infty} -\varphi'(t) \sin xt \, dt$$

Il est évident que $-\varphi'(t)$ est non-négative et non-décroissante, mais n'est pas une constante. Si $x > 0$, l'expression (2) peut s'écrire sous la forme :

$$\int_{0}^{\pi/x} \sum_{k=1}^{\infty} \left\{ \varphi'(t+\pi k/x) - \varphi'(t + \pi(k-1)/x) \right\} \sin xt \, dt$$

Puisque en chaque point de l'intervalle d'intégration, le k^e terme $\{\ \}$ de la somme Σ est non-négatif pour tout k, et positif pour certains k, et puisque $\sin xt$ est positif pour $0 < t < \pi/x$, il s'ensuit que la fonction (2) est positive pour tout $x > 0$. Cette remarque fondamentale fut faite, citant Steffensen [7], p. 184, par Lindhagen [4], puis par Polya ([5], p. 378, Théorème VII où, cependant, la restriction $f(\infty) = 0$, f étant $-\varphi'$, n'est pas mentionnée), et Steffensen [7], récemment, a amélioré ce point.

On voit aussi que la fonction impaire (2), outre qu'elle est positive pour tout $x > 0$, est continue en chaque point $x \neq 0$, puisque la convergence est uniforme ou voisine de chaque x, fixe, positive. Mais une intégration

par parties montre que le produit de x par la fonction (2) est identique à $\pi f(x)$, pour $-\infty < x < \infty$, $f(x)$ représentant la fonction

$$f(x) = \frac{1}{\pi}\int_0^\infty \varphi(t) \cos xt \, dt \tag{3}$$

(l'intégration par parties est légitime puisque $t\varphi'(t) \to 0$ que t tende vers 0 ou ∞, ainsi qu'il a été dit plus haut). Par conséquent l'hypothèse (1) une fois associée aux hypothèses : $\varphi(0) = 0$, $\varphi(\infty) = 0$ et $\varphi(t) = \varphi(-t)$, implique que $\varphi(t)$ est la transformée de Fourier-Stieltjes d'une fonction de distribution absolument continue ayant une densité $f(x)$ positive et continue pour $0 \leqq x < \infty$, et, par suite, également pour $-\infty < x < \infty$, puisque $f(x) = f(-x)$. Ce fait important et quelques-unes de ses conséquences ont été récemment soulignés par Polya [6], pp. 116-118. Pour une approche différente, cf. un article récent de Dugué et Girault [2].

Dans une note antérieure [9], pp. 592-594, ce théorème était combiné à un résultat de A. Kneser (cf. la présentation de Borel [1], pp. 46-52, et une correction de l'un des énoncés de Kneser dans [8], p. 95 ; la forme finale de l'énoncé correct découle de [10], pp. 601-602.

En vue de l'application donnée dans [9] aux équations différentielles, mais surtout en raison de l'application à la théorie des probabilités, ce qui va suivre répondra à la question : quand la densité $f(x) = f(-x)$ est-elle continue en $x = 0$?

Il est clair que la condition

$$\int^\infty \varphi(t) \, dt < \infty \qquad (\varphi \geqq 0) \tag{4}$$

suffit à cette fin, puisque (4) entraine la convergence uniforme de (3) pour $-\infty < x < \infty$. On montrera que cette condition suffisante est aussi nécessaire. Ceci est susceptible d'une interprétation simple aux termes de la considération géométrique de Dugué et Girault [2], pp. 3-5.

En fait, la démonstration sera faite de façon à prouver que si

$$\int^\infty \varphi(t) \, dt = \infty, \qquad (\varphi \geqq 0) \tag{4 bis}$$

alors non seulement $f(x)$ doit être discontinue au point $x = 0$ mais encore le devenir d'une façon tellement prononcée que

$$f(x) \to \infty, \qquad x \to 0 \tag{5}$$

En corollaire il s'ensuit que $f(x)$ doit être continue au point $x = 0$ chaque fois que

$$f(x) < \text{constante}, \quad x \to 0 \tag{5 bis}$$

Un autre corollaire sera que ou bien (5) est vrai ou bien $f(x)$ est non seulement continue pour $-\infty < x < \infty$ mais en plus $f(x) \to 0$ quand $|x| \to \infty$.

En fait cette dernière relation découle de (3) en vertu du lemme de Riemann-Lebesgue, si elle est applicable c'est-à-dire si (4) est satisfait.

Soit $g(t)$ une quelconque fonction définie sur le demi-droit ouvert $0 < t < \infty$ de façon à remplir les conditions suivantes : $g(t)$ est non-négative et monotone, $g(\infty) = 0$, et

(6a) $$\int_{+0}^{1} t g(t)\, dt < \infty ,$$ (6b) $$\int_{1}^{\infty} g(t)\, dt = \infty .$$

Alors, ainsi que (3) le montre, la fonction

(7) $$h(x) = \int_{0}^{\infty} g(t) \sin xt \, dt$$

doit satisfaire à la relation

(8) $$h(x)/x \to \infty , \quad x \to 0$$

Choisissons

(9) $$g(t) = -\varphi'(t) \ , \text{ où } \ 0 < t < \infty .$$

Alors il est clair d'après les remarques faites au début de la note que les conditions spécifiées pour $g(t)$ avant (6a) et (6b) sont remplies, de même que l'est la condition (6a) (et ceci, que l'on soit dans le cas (4) ou (4 bis)). Mais si on suppose (4 bis), alors la condition (6b) est aussi remplie. De fait, (6b) équivaut à (4 bis) en vertu de (9). Par conséquent, (8) est vrai dans le cas présent.

D'autre part, on voit d'après (8) et (9) que $h(x)$ est précisément l'intégrale (2). Puisque l'intégration par parties, faite plus haut, montre que le produit de πx par la fonction (3) est identique à l'intégrale (2), il s'ensuit que $h(x) = \pi x f(x)$. Donc, (8) prouve (5).

The Johns Hopkins University

BIBLIOGRAPHIE

[1] E. Borel, Leçons sur les séries divergentes, Paris, 1901.

[2] D. Dugué et M. Girault, "Fonctions convexes de Polya", Publications de l'Institut de Statistique de l'Université de Paris, vol. 4 (1955), pp. 3-10.

[3] P. Hartman and A. Wintner, "On the behavior of Fourier sine transforms near the origin", Proceedings of the American Mathematical Society, vol. 2 (1951), pp. 398-400.

[4] A. Lindhagen, Studier ofver Gamma-Funktionen och nagra beslagtade Transcendenter, Uppsala, 1887.

[5] G. Polya, "Ueber die Nullstellen gewisser ganzer Funktionen", Mathematische Zeitschrift, vol. 2 (1918), pp. 352-383.

[6] _______, "Remarks on characteristic functions", Berkeley Symposium on Mathematical Statistics and Probability, (1949), pp. 115-123.

[7] J. F. Steffensen, " Bounds of certain trigonometrical integrals", Tenth Scandinavian Congress of Mathematicians, Copenhagen, 1947, pp. 181-186.

[8] A. Wintner, "On the Laplace-Fourier transcendents occuring in mathematical physics", American Journal of Mathematics, vol. 69, (1947), pp. 87-98.

[9] _______, "A priori Laplace transformations of linear differential equations", American Journal of Mathematics, vol. 71 (1949), pp. 587-594.

[10] _______, "On almost free linear motions", ibid., vol 71 (1949), pp. 595-602.

ANALYSE D'ARTICLES PARUS DANS DES REVUES ÉTRANGÈRES

BULLETIN OF MATHEMATICAL STATISTICS

Vol. 4 - N°s 1-2 Décembre 1950

Kyushu University, *Fukuoka - JAPAN*

THE DISTRIBUTION OF STATISTICS DRAWN FROM THE GRAM-CHARLIER TYPE A POPULATION, par URANISI

Si M et σ sont la moyenne et l'écart-type de la distribution d'une variable aléatoire, l'auteur assure pour cette distribution un développement de Gram-Charlier de type A jusqu'au 5e ordre et calcule les distributions d'échantillonnage du statistique :

$$t = \frac{n\,(\bar{x} - M)}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}} \qquad (\bar{x} = \text{moyenne empirique calculée à partir de } x_1, \ldots, x_n)$$

puis du rapport F de deux variances empiriques indépendantes. Il en conclut quelques résultats concernant l'effet de la non normalité sur les distributions de t et du F classiques.

RANDOM INTEGRATIONS, par T. KITAGAWA

a) f(t) est une fonction définie et continue dans l'intervalle fermé (0, 1) une approximation de l'intégrale de f(t) sur (0, 1) est de la forme :

$$S_A(f) = \sum_{i=1}^{n} A_i f(t_i) \qquad 0 \leq t_1 < t_2 < \ldots < t_n \leq 1$$

$$\sum_{i=1}^{n} A_i = 1 \qquad A_i > 0$$

L'auteur envisage trois types d'intégration A, B, C : (A) f(t) est une fonction fixée, les points $t_1 \ldots t_n$ et les A_i sont sélectionnés au hasard.

b) f(t) appartient à une famille donnée de fonctions avec une probabilité convenablement définie (mesure de Wiener) et S_A est une variable aléatoire ; (C) est une combinaison de (A) et (B).

Pour chaque type le mode de détermination des t_i et A_i peut être envisagé sous 3 aspects. Alors l'auteur étudie les moyenne et variance des différentes sommes $S_A(f)$ ainsi définies.

NOTE OF THE USE ORDER STATISTICS, par K. KONO

Le but de cet article est de fournir un procédé d'échantillonnage permettant une estimation suffisamment précise de la moyenne d'une variable aléatoire x au moyen d'une variable aléatoire auxiliaire y corrélée avec x pour laquelle l'échantillonnage est meilleur marché. La distribution de (x,y) est symétrique autour de l'origine (0,0) et la régression de x en y est linéaire. Lorsque la distribution de (x,y) est normale, il est possible de déterminer un procédé optimum (variance minimum de l'estimateur) qui assure un coût total d'échantillonnage C fixé.

NOTE ON DOUBLE SAMPLING METHOD, par K. KONO

Le problème traité est celui de trouver une condition pour laquelle une estimation basée sur la méthode d'échantillonnage double stratifié peut être plus précise que celle d'une méthode d'échantillonnage simple pour un coût total de l'échantillonnage fixé.

SOME CONSIDERATIONS ON THE RATIC AND REGRESSION ESTIMATES par M. MURAKAMI

L'objet de cette note est de comparer les efficacités de l'estimateur qui consiste à utiliser le rapport $\frac{\overline{x}}{\overline{y}}$ des moyennes empiriques des variables aléatoires x et y corrélées (x variable auxiliaire) et la régression avec la moyenne empirique $\overline{y}$ pour estimer la moyenne m_y de la variable y en faisant intervenir des considérations du coût d'échantillonnage.

Vol. 5 - N°s 1-2 Septembre 1952

NOTE ON ENUMERATION OF 7 x 7 LATIN SQUARES, par K. YAMAMOTO, p. 1-9

L'auteur reprend la méthode de Sade et étend ce mode de classification à tous les carrés latins en mettant en évidence une relation symétrique entre matrices incidentes fondamentales.

STOCHASTICS STUDIES ON THE ATOMIC BOMB CASUALTIES, par M. MASUYAMA p. 21-31

L'auteur montre comment l'utilisation de quelques méthodes statistiques d'échantillonnages a permis de diriger une étude relative au nombre des décès causés par la bombe atomique à Hiroshima.

ON THE STATISTICAL INFERENCES IN FINITE POPULATIONS BY TWO SAMPLE THEORY, par H. URANISI, p. 9-21

Cet article fait suite à l'exposé de T. KITAGAWA (Mai 1951) au meeting annuel de la Société mathématique du Japon où il établit une méthode d'inférences stochastiques sur les populations finies en considérant celles-ci comme un échan-

tillon de grandeur N d'une population infinie π. En supposant cette population infinie π normale, l'auteur fournit des intervalles de confiance avec un degré de signification $(1-\alpha)$

1°) pour la moyenne x et la variance S^2 de la population finie correspondante.
2°) pour la différence $x_1 - x_2$ entre les moyennes de deux populations finies ainsi que pour le rapport :

$\frac{S_2^2}{S_1^2}$ des variances de ces deux populations. Utilisant ensuite une approximation de Gram-Charlier de type A, il montre les effets de la non-normalité de π sur les distributions de la moyenne $\overline{x}$ et de la variance, s^2, empiriques de la population finie.

CONVERGENCE OF INTEGRAL AND ITS APPLICATIONS, par M. IZAKI, p. 31-35

Si $f(t)$ est une fonction continue détinie dans $0 \leq t \leq \ell$; la somme:

$S_{(1)} = \sum_{k=1}^{n} f(\xi_k)(t_k - t_{k-1})$ converge lorsque $\text{Max}(t_k - t_{k-1}) \longrightarrow 0$. où $t_0 = o$; $t_n = \ell$

Si les $n \neq$ points $t_1, \ldots t_{n-1}$ sont triés au hasard et les ξ_k choisis au hasard dans (t_{k-1}, t_k), $S_{(1)}$ est une variable aléatoire qui converge en probabilité vers $\int_0^{\ell} f(t)\, dt$ quand $n \longrightarrow \infty$;

Lorsque les t_k ne sont plus aléatoires, mais l'intervalle $(0, \ell)$ divisé en n strata de longueur $\frac{1}{n}$, les résultats sont analogues. L'auteur montre une application à un intégrateur pour lequel le calcul est fait de telle manière que t soit prélevé séquentiellement et t soit un processus aléatoire de Poisson.

SUCCESSIVE PROCESS OF STATISTICAL INFERENCES, par T. KITAGAWA, p. 35-51

Dans une première partie, l'auteur montre comment substituer à l'écart-type empirique lors de l'utilisation du test du t avec une application aux cartes de contrôle ; en particulier si $\overline{x}, \ldots \overline{x}_{n-1}, \omega_1, \omega_2, \ldots \omega_{n-1}$, sont les moyennes et étendues empiriques calculées à partir d'échantillons $0_1, 0_2, \ldots 0_{n-1}$ provenant d'une même population, il est possible de déterminer un test ayant un degré de signification $(1-\alpha)$ pour étudier si un échantillon O_n provient de cette même population.

Dans une deuxième partie, l'auteur obtient pour le test de Fisher-Behrens (comparaison de deux populations normales de moyennes et variances inconnues) une région critique dont les limites sont fixées à l'avance ; il en déduit un intervalle de confiance pour la différence entre les moyennes de deux populations normales d'étendue égale à ℓ fixée.

RANDOM FREQUENCY PROCESS, par T. ONOYAMA, p. 51-53

Soit $x(t) = \sum_{n=1}^{\infty} a_n(\omega)\, e^{i\lambda_n(\omega)t}$ un processus stochastique où les a_n, λ_n sont des variables aléatoires réelles, indépendantes telles que

$$\sum_{n=1}^{\infty} E(a_n^2) \leq \infty$$

Un tel processus est dit (r.f.p.) (random frequency process). L'auteur montre: 1°) qu'un processus r.f.p. est stationnaire au sens faible et il est possible de

construire un processus r.f.p. ayant une distribution spectrale continue ou une fonction d'autocorrélation donnée de type stationnaire.

2°) Le spectre $S(\lambda,\omega)$ d'un processus r.f.p. est $S(\lambda) = \sum_{n=1}^{\infty} a_n \psi_n(\lambda)$

où ψ_n est une fonction aléatoire indépendante des a_n ne pouvant prendre que le système de valeurs $(0,1)$; $(-\frac{1}{2}, +\frac{1}{2})$; $(-1,0)$ et réciproquement.

MINIMAX ESTIMATIONS, par K. MIYASAWA, p. 53

Cet article fait suite à celui "On the statistical decision function I". Bull. Math. Stat. Research, vol. 4 - 1950.

Vol. 5 - Nos 3-4 Juin 1953

ON THE FILTER PROBLEM OF A STATIONARY STOCHASTIC PROCESS, par S. KANO, p. 47

Etant donné un processus stochastique stationnaire $x(t)$ représentant un message, et deux processus stationnaires $y(t)$ et $z(t)$ représentant le bruit, il est possible de déterminer une fonction $K_{(\sigma)}$ σ qui minimise :

$$E\left[\left|x(t+\alpha) - \int_{-\infty}^{+\infty}\left\{y(t-\sigma) + x(t-\sigma)\, z(t-\sigma)\right\} dK(\sigma)\right|^2\right]$$

$$\text{Si} \quad k(\omega) = \int_0^{\infty} e^{-i\omega\sigma}\, dK(\sigma)$$

$k(\omega)$ est une fonction rationnelle, n'ayant pas de zéros dans le 1/2 plan inférieur tandis que le numérateur n'a aucune singularité dans ce même 1/2 plan.

NOTE SUR L'ESTIMATION DE LA VALEUR MOYENNE D'UN PROCESSUS STOCHASTIQUE, par A. KUDO, p. 53-58

Le but de cet article est de discuter quelques combinaisons de la théorie de l'estimation linéaire de la fonction valeur moyenne d'un processus stochastique dues à O. GRENENDER et la notion d'intégration aléatoire de type A introduite par T. KITAGAWA. Dans le deuxième paragraphe est étudiée l'estimation de la fonction valeur moyenne d'un processus stochastique sous la condition que l'observation continue de la réalisation du processus stochastique n'est pas possible. Le troisième paragraphe est consacré à la discussion du cas spécial où le processus est stationnaire.

SOME STOCHASTIC CONSIDERATIONS UPON EMPIRIAL FUNCTIONS OF VARIOUS TYPES, par T. KITAGAWA, p. 19

L'auteur expose quatre procédés pour construire un modèle mathématique représentatif d'un phénomène dans un intervalle de temps $(0,1)$. Les cas suivants sont envisgés :

(A) Il existe une fonction certaine g(t) inconnue, $\in L^2(0,\infty)$;

(B) La fonction g(t) est un processus stochastique de la forme $\sum_{\nu=1}^{\infty} \lambda_\nu^{-\frac{1}{2}} \varphi_\nu(t)\, z_\nu$ où les $\varphi_\nu(t)$ et λ_ν sont connues et les z_ν sont des variables aléatoires normales, indépendantes (a_ν, σ^2).

Il s'agit de représenter g(t) à l'aide d'une fonction empirique $\hat{g}(t,n)$ formée :
(D) à partir des résultats des r_i observations $y_{ij} = g(t_i) + \varepsilon_{ij}$. Aux points t_i fixés à l'avance [les ε_{ij} sont des variables aléatoires normales indépendantes $(0, \sigma^2)$];
(C) à partir des n réalisations $f_i(t)$ de g(t) dans un même intervalle de temps.

Dans une deuxième partie, l'auteur donne une approximation empirique de la dérivée de g(t) mais signale des difficultés rencontrées dans ce problème.

ON THE DETERMINATION OF SAMPLE SIZE FROM THE TWO SAMPLES THEORICAL FORMULATION, par T. KITAWAGA, T. KITAHARA, NOMACHI ; N. WATANABE, p. 35

Le problème est de déterminer pour la moyenne m d'une population normale d'écart-type σ inconnu, un intervalle de confiance (niveau $1-\alpha$) dont l'étendue est inférieure à 2α fixée. Un premier échantillon $O_{n_1}(x_{11}, x_{12}, \ldots x_{1n_1})$ est prélevé à l'aide duquel est calculée la variance empirique s_1 ;

Une formule permet de déterminer la taille n_2 d'un deuxième échantillon à partir de s_1, de manière qu'il y ait une probabilité supérieure à $(1-\beta)$ que la 1/2 étendue de l'intervalle de confiance pour m, évalué à l'aide de ce deuxième échantillon soit inférieur à α.

Vol. 6 - N°s 1-2 Décembre 1955

A CONTRIBUTION TO A NOTATION SYSTEM OF THE CONFUNDED FACTORIAL EXPERIMENTS, par T. KITAGAWA et M. MITOME, p. 1-11

Le but de cet article est d'expliquer les nouveaux systèmes de notation proposés par les auteurs à "the branch Meeting of Japonese Math. Soc. "Kumamoto University fev. 1951".

ON THE WEIGHTED POWER FUNCTION OF SOME TESTING HYPOTHESES, p.11-17

Soit $\theta\,(\theta_1, \ldots \theta_s)$ un paramètre dont le domaine des valeurs possibles est Ω ; ω_0, ω_1 deux suites disjointes de Ω, pour tester $(H_0)\ \theta \in \omega_0$ contre $H_2\ \theta \in \omega_1$, l'auteur détermine une région critique qui minimise l'aire limitée par la fonction puissance du test dans le cas où $\theta \in \omega_1$.

Si ω_0 se réduit à $\theta = \theta_0$ et H_1 est $\theta \in \omega_1$, il retrouve le lemme de Neymann Pearson ; si ω_0 est quelconque, $\mu(\theta)$ est la distribution de θ dans ω_0, $\varphi_0(x)$ la probabilité d'accepter H_0 quand x est la valeur observée de $X(X_1 \ldots X_n)$

Alors $\varphi_0(x)$ est définie par

$$\varphi_0(x) = 1 \quad \text{quand} \int_{\omega_1} f(x,\theta)\, d\theta \geqslant k \int_{\omega_0} f(x,\theta)\, d\mu(\theta)$$
$$\varphi_0(x) = 0 \quad \text{quand} \int_{\omega_1} f(x,\theta)\, d\theta < k \int_{\omega_0} f(x,\theta)\, d\mu(\theta)$$

où k est une constante.

De plus $\varphi_0(x) \in \Phi_\alpha$, classe des tests φ pour lesquels la puissance $\beta(\theta,\varphi) \leqslant \alpha$ pour tout $\theta \in \omega_0$ et $\beta(u,\varphi_0) = \int_{\omega_0} \beta(\theta,\varphi_0)\, d\mu(\theta) = \alpha$

TABLES FOR TESTING RANDOMNESS BY MEANS OF LENGHTS OF RUNS, par M. TAKASHIMA, p. 17-25

Si $Q_1(t)$ est la probabilité d'apparition d'au moins une suite de A de longueur t ou plus, $Q(t)$ celle d'au moins une suite de A ou B de longueur t ou plus, l'auteur désigne par t_5, t_1, $\tilde{t}_5$, $\tilde{t}_1$ les plus petits entiers t pour lesquels $Q(t)$ $[Q_1(t)]$ est inférieur à 0,05 ou 0,01. En désignant par m et n le nombre total des A et des B, F. MORTELLER avait déjà calculé ces valeurs lorsque m = n = 5 ; 10 ; 15 ; 20 ; 25.

[Note on an application of runs to quality control charts, Ann. Math. Statist. (12) - 1941 -]

Dans cette note quatre tables fournissent les valeurs de t_1, t_5, $\tilde{t}_1$, $\tilde{t}_5$ pour toutes les combinaisons de m et n avec $\begin{matrix} 1 \leqslant m \leqslant 25 \\ 1 \leqslant n \leqslant 25 \end{matrix}$

AUXILIARY TABLES FOR THE APPLICATIONS OF n-DIMENSIONAL t. DISTRIBUTIONS TO CERTAIN CLASS OR EMPIRICAL FUNCTIONS, par Y. NOMACHI, p. 25-49

Cet article fait suite à celui de T. KITAGAWA : "Some stochastic Considerations upon empirical Functions" (Bull. Math. Statist. Col. 5 - n° 3-4) et fournit des tables permettant de calculer rapidement les fonctions empiriques proposées par T. KITAGAWA pour représenter une fonction inconnue g(t).

MITTEILUNGSBLATT FUR MATHEMATISCHE STATISTIK
Jahrgang 7 - 1955 - Heft 3

LES MODÈLES STATISTIQUES COMME AUXILIAIRES DANS LA DESCRIPTION DE LA NATURE, par H. VON SCHILLING

L'auteur décrit quelques modèles statistiques qui lui paraissent plus aptes à décrire la réalité :

1°) On étudie un évènement ayant les propriétés suivantes : a) il se produit avec la probabilité p par unité de temps, b) une fois qu'il s'est produit, il ne peut se produire dans les k unités de temps suivantes, c) il a été observé au temps n.

On calcule la probabilité W(m,n,k,p) que cet évènement ait été le m^{ième} depuis le temps O. Le modèle ainsi défini fournit une généralisation de celui de Bernoulli, car pour k = 0 on obtient la distribution de Bernoulli B(m-1,n-1,p).

Le modèle parait apte à décrire des phénomènes nécessitant un temps de récupération (utilisation dans la statistique des caisses de maladie).

2°) Soit X_i une suite de v.a. laplaciennes réduites. On sait que $S_n = \sum_{i=1}^{n} X_i$ est une v.a. laplacienne d'e.m. nulle et dont l'e.q.m. $\sigma_n = \sqrt{n}$ tend vers ∞ lorsque $n \to \infty$. Il existe des phénomènes presque périodiques où un tel accroissement de σ_n

est impossible. Soit par exemple V_i le volume d'air inspiré par une personne en temps i. Le volume est soumis à des fluctuations et si l'on pose $D_1 = V_1 - V_0$, $D_2 = V_2 - V_1$ ..., $D_n = V_n - V_{n-1}$, il vient $V_n = V_0 + \sum_{i=1}^{n} D_i$. Il est évident que D_i ne peut suivre une loi de Laplace, sans quoi V_n tendrait vers l'∞ lorsque $n \to \infty$. L'auteur propose pour D_i la distribution de densité :

$$w(y) = \frac{\int_y^{\infty} e^{-\frac{x^2}{2\sigma^2}} dx}{\int_0^{\infty} x\, e^{-\frac{x^2}{2\sigma^2}} dx}$$

3°) On définit des cheminements aléatoires sur une surface déterminée de la façon suivante : une particule part du point O sur la surface suivant un élément de courbe possédant en O la courbure géodésique Kg(O). Après avoir parcouru un élément de courbe de longueur constante ds, elle peut changer de direction suivant une loi normale. On cherche les chemins les plus probables. On montre qu'ils sont définis par la condition suivante : $\int_0^{\infty} Kg^2(s)\, ds$ = minimum. L'auteur donne des solutions explicites pour le plan et la surface de la sphère.

4°) Pour décrire, en mécanique, le passage du repos au mouvement, l'équation $\vec{F} = m\vec{\Gamma}$ ne suffit pas ; celle-ci n'est valable qu'après l'établissement d'un certain régime précédé de phénomènes transitoires. L'auteur propose un modèle statistique pour décrire ces phénomènes transitoires, modèle qui conduit à postuler l'existence d'un intervalle de temps élémentaire.

LA DISTRIBUTION NORMALE TRONQUÉE TRAITÉE GRAPHIQUEMENT SUR DU PAPIER DE PROBABILITÉ, par K. STANGE

Un procédé de fabrication a fourni un lot d'objets dont la caractéristique suit une loi normale (μ, σ). Si la frontière de tolérance T est située à droite de $\mu - 3\sigma$, il faut rejeter du lot les objets défectueux dont le pourcentage sera désigné par A^*. Le pourcentage des objets prêts pour la consommation sera alors nécessairement $B^* = 1 - A^*$.

Soit F (x) la fonction de répartition de X. On a alors évidemment $F(T) = A^*$. Etudions l'ensemble B^* séparément. B^* est caractérisé au moyen des paramètres μ, σ, T et de $A^* = F(T)$. Le problème se pose alors de trouver, à partir de la fonction de répartition empirique B (x) de B^*, des estimations $\hat{T}, \bar{x}, s, \hat{A}$ pour les paramètres T, μ, σ, A^*. L'auteur donne une solution graphique de ce problème en employant du papier de probabilité. Application à un problème concret.

UN NOUVEAU GENRE DE NOMBRES ; SES PROPRIÉTÉS ET SES APPLICATIONS EN STATISTIQUE MATHÉMATIQUE, par I. LAH

Posons $(X)_n = X(X+1)(X+2)\ldots(X-n+1)$. Les nombres de Stirling de 1ère espèce S_n^v sont définis par l'identité $(X)_n = \sum_{v=1}^{n} S_n^v X^v$ et ceux de 2e espèce Υ_n^v par l'identité $X^n = \sum_{v=1}^{n} \Upsilon_n^v (X)_n$.

L'auteur définit un nouveau genre de nombres par analogie de la manière suivante: Posons $[X]_n = X(X+1)(X+2)\ldots(X+n-1)$. On définit des nombres s_n^v par l'identité $(X)_n = \sum_{v=1}^{n} (-1)^v\, s_n^v (X)_v$ ou l'identité $(X)_n = \sum_{v=1}^{n} (-1)^v\, s_n^v (X)_v$. Ces nombres ont des propriétés remarquables très faciles à démontrer. On peut par exemple les représenter en fonction des nombres de Stirling de 1ère et de 2e espèce suivant: $s_n^v = \sum_{i=v}^{n} (-1)^i\, S_n^i\, \Upsilon_i^v$.

Comme applications à la statistique, notons les identités :

$$E\{[\chi]_n\} = \sum f(\chi)\,[\chi]_n = \sum_{\nu=1}^{n} (-1)^{\nu}\, s_n^{\nu}\, E\{(\chi)_n\}$$

$$E\{(\chi)_n\} = \sum f(\chi)\,(\chi)_n = \sum_{\nu=1}^{n} (-1)^{\nu}\, s_n^{\nu}\, E\{[\chi]\}$$

ainsi que le résultat suivant : Définissons un opérateur θ par $\theta = \chi^2 \frac{d}{d\chi}$. Alors $\left(\frac{d}{d\chi}\right)^n = \sum_{\nu=1}^{n} (-1)^{\nu} s_n^{\nu}\, \chi^{-n-\nu}\, \theta^{\nu}$. A partir de cette formule l'auteur établit quelques identités concernant les fonctions génératrices.

SUR LA SURFACE INTERIEURE D'UN AMAS DE SPHERES, par H.L. DE VRIES

Considérons une distribution au hasard de sphères dans l'espace. On s'intéresse souvent à la densité de probabilité $f(x)$ de la distribution des diamètres x de ces sphères. Dans le présent travail l'auteur se propose de déterminer la surface totale O_i des sphères contenues dans un volume V. A cet effet on n'a besoin que du 2e, éventuellement du 3e moment de $f(x)$. Comme application il faut songer à la surface intérieure O_i d'un volume V de coke. La connaissance de O_i est importante car l'activité du coke dans le haut-fourneau lui est liée.

Jahrgang 8 - 1956 - Heft 1

PASSAGE DE LA SOLUTION DU CENTRE DE GRAVITÉ A CELLE DES AXES PRINCIPAUX ET AUTRES DEVELOPPEMENTS DES MÉTHODES D'APPROXIMATION EN ANALYSE FACTORIELLE, par Von Rolf Bargmann

En analyse factorielle la méthode des moindres carrés conduit à la solution dite des axes principaux (principal axes solution) qui a l'inconvénient de mener à des calculs fastidieux. Le présent travail expose une méthode qui, partant de la solution habituelle du centre de gravité, permet un passage rapide à celle des axes principaux. Nombreuses applications.

THÉORÈMES LIMITE RELATIFS A DES PROBABILITÉS DE PASSAGE TENDANT VERS O, par V. KANNGIESSER

On considère une chaine de Markoff discrète admettant deux états possibles 0 et 1. Soit $V = \begin{pmatrix} v_{00} & v_{01} \\ v_{10} & v_{11} \end{pmatrix}$ la matrice des probabilités de passage. On suppose que l'on est dans le cas régulier. Deux problèmes se posent :

1°) Problème de Bernoulli : déterminer la probabilité $w_n(x,V)$, $x = 1,\ldots,n$ pour que, dans les n premières épreuves il y ait x états 1.

2°) Problème de Pascal : Déterminer la probabilité $\hat{w}_m(z,V)$, $z = 0,1,\ldots$ pour que le $m$ième état 1 se produise à la $m+z$ ième épreuve.

L'auteur se propose d'étudier le comportement asymptotique des probabilités $w_n(x,V)$ et $\hat{w}_m(z,V)$ lorsque n (respectivement m) tendent vers l'infini. Il montre en

particulier que si l'une des probabilités de passage v_{ij} ; i,j = 0,1 tend vers 0 comme $\frac{1}{n}$ les quantités $w_n(x,V)$ et $\hat{w}_n(x,V)$ tendent vers une limite commune qui n'est autre que la distribution de Poisson $\psi(x,k) = \frac{k^x}{x} e^{-k}$. Un autre cas de limite commune est discuté.

PROCÉDÉS DE CONTRÔLE POUR DES VARIABLES A INTERVALLE DE CONFIANCE LARGE ET ETROIT, par **K.** BRUCKER-STEINKUHL

Travail d'exposé illustré par de nombreux exemples.

SUR LA THÉORIE STOCHASTIQUE DE L'ITÉRATION DES CARACTÈRES (RUNS) par Otto LUDVIG

L'auteur donne une excellente introduction historique, un aperçu pratiquement complet de l'état actuel de la théorie des runs ainsi qu'une bibliographie exhaustive. Il se pose ensuite le problème suivant : Déterminer la probabilité pour que, dans une suite ordonnée de n éléments il y ait exactement (ou au moins) ν itérations. Ce problème peut être résolu de deux manières différentes, soit par l'analyse combinative, soit par la théorie des équations aux différences finies. C'est à cette dernière méthode que l'auteur donne sa préférence.

APPLICATION DE LA LOI LOG-NORMALE, par Rolf WARTMANN

Suite d'un travail antérieur paru dans cette revue 7 (1955), sept. 2 . Dans ce travail antérieur l'auteur avait montré comment on peut passer d'une distribution log-normale à une distribution normale empirique. Il avait également établi des formules pour les variances des paramètres empiriques d'une telle distribution log-normale, à savoir : moyenne μ, e.m.q. σ, asymétrie γ_1.

Dans le présent travail, l'auteur déduit la variance du point de fuite empirique a. Il donne également des indications relatives à l'étude de l'échantillonnage lorsqu'on a affaire à une distribution log-normale. Application à un problème concret.

J. & R. SENNAC, Imprimeurs
54, Fg Montmartre, PARIS (9e)